DOCUMENT RESUME

ED 249 292 TM 840 743

AUTHOR	Fetters, William B.; And Others
TITLE	Quality of Responses of High School Students to Questionnaire Items. High School and Beyond: A National Longitudinal Study for the 1980's.
INSTITUTION	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	NCES-84-216
PUB DATE	Sep 84
NOTE	64p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Grades (Scholastic); High Schools; *High School Students; Longitudinal Studies; National Surveys; Parents; *Questionnaires; *Reliability; Student Characteristics; Twins; *Validity
IDENTIFIERS	*High School and Beyond (NCES)

ABSTRACT

Three features of High School and Beyond (HS&B) provide rare opportunities for examining the quality of the responses of high school students to group-administered questionnaires: (1) many identical or similar items also were asked in questionnaires completed by about 6,500 of the students' parents; (2) high school transcripts containing information about coursework and grades were obtained in fall 1982 for about 16,000 of the 1980 sophomore cohort HS&B participants; and (3) questionnaire data were obtained from both members of more than 500 twin pairs. This study evaluated the validity of student responses, using the parent questionnaire data as a standard; the validity of student reports of their grades, using transcript data as a standard; and the reliability of student responses, estimated from twin data. The results of analyses of the three HS&B data sets were generally quite consistent with the findings of previous investigations. The quality of student questionnaire data depended on both the nature of the questions asked and the characteristics of the student who provided the answers. (Author/BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED249292

High School and Beyond

a national longitudinal study for the 1980's

Quality of Responses of High School Students to Questionnaire Items

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. Kirschhammer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

National Center for Education Statistics

Tm 84s 743

High School and Beyond

a national longitudinal study for the 1980's

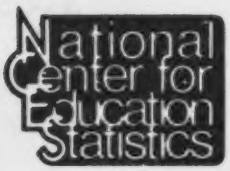
Quality of Responses of High School Students to Questionnaire Items

William B. Fetters
Peter S. Stowe
Jeffrey A. Owings

U.S. Department of Education
T. H. Bell
Secretary

Office of Educational Research and Improvement
Donald J. Senese
Assistant Secretary

National Center for Education Statistics
Emerson J. Elliott
Administrator



National Center for Education Statistics

"The purpose of the Center shall be to collect and disseminate statistics and other data related to education in the United States and in other nations. The Center shall . . . collect, collate, and, from time to time, report full and complete statistics on the conditions of education in the United States; conduct and publish reports on specialized analyses of the meaning and significance of such statistics; . . . and review and report on education activities in foreign countries."--Section 406(b) of the General Education Provisions Act, as amended (20 U.S.C. 1221e-1).

September 1984

Table of Contents

	Page
Foreword	v
Executive Summary	vii
Chapter I. Introduction	1
1.1 Background and Purpose	1
1.2 Findings of Previous Studies	2
1.2.1 Stability and Validity Study, 1972 Seniors	2
1.2.2 Reliability of Retrospective Data	3
1.2.3 Validity and Reliability of Survey Research Questionnaires	3
1.3 Sources of Data	4
1.3.1 Twin Component of Student Survey	4
1.3.2 Parent Survey	4
1.3.3 Transcript Study	5
Chapter II. Methodology	7
2.1 Variables Analyzed	7
2.2 Treatment of Data	7
2.3 Sample Sizes	8
2.4 Subgroup Comparisons	10
2.5 Measurement of Validity, Bias, Reliability, and Consistency	11
Chapter III. Findings	13
3.1 Family Background Variables	13
3.1.1 Validity—Parent as Standard	13
3.1.2 Bias—Parent as Standard	18
3.1.3 Reliability—Twin File	18
3.1.4 Comparison of Validity and Reliability Coefficients	19
3.2 School-Related Variables	19
3.2.1 Validity—Transcript as Standard	19
3.2.2 Bias—Transcript as Standard	23
3.2.3 Reliability—Twin File	24
3.3 Other Variables	25
3.3.1 Consistency of Parent-Child Responses—Grades, Plans, Attitudes	25
3.3.2 Mean Difference Between Parent and Child Responses—Grades, Plans, Attitudes	27
3.3.3 Consistency of Parent-Child Responses—Financial and College Items	27
3.3.4 Mean Difference Between Parent and Child Responses—Financial and College Items	29
Chapter IV. Summary and Discussion	31
A. Data Collection	31
B. Composites and Means	31
C. Statistical Adjustment and Data Interpretation	32
D. Future Surveys	32
References	33
Appendix	37
Estimated Coefficients of Validity, Bias, Reliability and Consistency by Item Type for Selected Student Subgroups: 1980 High School Sophomores and Seniors	

Foreword

High School and Beyond (HS&B) is a large-scale, national longitudinal study designed and sponsored by the National Center for Education Statistics (NCES), with support from other governmental agencies. HS&B provides a variety of data about 1980 sophomores and seniors as they move through U.S. high schools and into the many and varied activities of early adulthood. The study began with the group administration of questionnaires and tests to 30,000 sophomores and 28,000 seniors, including over 500 sets of twins, in more than 1,000 public and private schools in spring 1980. In fall 1980, data were collected from over 6,500 of the parents of HS&B participants, primarily to investigate questions regarding financial planning for postsecondary education. HS&B has continued with a second collection of information from the 1980 sophomores and seniors in spring 1982 and the collection of high school transcripts in fall 1982 for a subsample of sophomore cohort members. A third data collection from 1980 sophomores and seniors took place in spring 1984.

It is not possible to obtain 100 percent cooperation in a strictly voluntary survey such as HS&B, and those who do respond may not always possess accurate information or for other reasons may provide inaccurate answers. Good survey practices require the examination of the quality of the data collected. Assessment of data quality leads to better analysis and interpretation of the data and improvements in the designs of future studies. Mechanisms for examining data quality were built into the HS&B design. Nonresponse bias is considered elsewhere. This report examines the quality of responses of high school students to questionnaires that were group administered in a school setting. The validity of student responses, in most instances, is judged against the standard of parent responses; and reliability is estimated by comparing the responses of the two members of twin data. The validity of student reports of grade averages and courses taken is assessed by use of transcript data. Data quality is evaluated as a function of item type and the age, sex, race/ethnicity, and other characteristics of the respondents.

The results of this study will be useful in the analysis and interpretation of the data generated by HS&B and similar surveys, and will be helpful in designing student questionnaires for future surveys.

David A. Sweet, Assistant Administrator
Division of Multilevel Education Statistics

C. Dennis Carroll, Chief
Longitudinal Studies Branch

Executive Summary

Three features of High School and Beyond (HS&B) provide rare opportunities for examining the quality of the responses of high school students to group-administered questionnaires: (1) many identical or similar items also were asked in questionnaires completed by about 6,500 of the students' parents; (2) high school transcripts containing information about coursework and grades were obtained in fall 1982 for about 16,000 of the 1980 sophomore cohort HS&B participants; and (3) questionnaire data were obtained from both members of more than 500 twin pairs. The findings of previous investigations of the reliability or validity of survey data obtained from high school students are discussed in the report. The results of analyses of the three HS&B data sets are generally quite consistent with the findings of previous investigations. As indicated below, the quality of student questionnaire data depends on both the nature of the questions asked and the characteristics of the student who provides the answers.

A lag of about 6 months between the collection of data from parents and students and other considerations (e.g., inaccuracies in responses of parents) make the standards against which accuracy is judged somewhat fallible. Thus, actual quality of student questionnaire data undoubtedly is somewhat higher in many instances than that indicated in this report.

A. Validity—Parent Questionnaire Data as Standard.

1. The quality of HS&B student questionnaire data generally is fairly high for contemporaneous, factual information. For example, the validity coefficients* are almost .90 for father's educational attainment and about .60 for father's occupation. The validity of the socioeconomic status (SES) composite that often is employed in HS&B data analyses was found to be at least .80.
2. The quality of retrospective information tends to decline with the passage of time; for example, concerning whether the mother worked at various times, from about .71 for during high school, to about .64 for during elementary school, to .53 for prior to elementary school.
3. The validity of less factual information tends to be lower; e.g., about .60 for the mother's aspirations for her child's education.

B. Validity—Transcript Data as Standard

1. The grade averages reported by seniors correlate well (.77) with grade point averages computed from their transcripts. Seniors reported their grades to be somewhat higher (about one-fourth of a letter), however, than shown by their transcripts.
2. The correlation coefficients between senior reports and transcript data are high (in the .80's) for amount of coursework in specific foreign languages and for whether geometry, physics, or chemistry ever was taken. The coefficients are somewhat lower, ranging from .63 to .70, for amounts of coursework in mathematics and science and for whether 2nd-year algebra, trigonometry, and calculus ever were taken. They are lower yet for the two areas (history or social studies and English or literature) that show relatively little variation from student to student in amount of coursework taken. Seniors tended to report they had taken more coursework in most areas than reflected by their transcripts. The amount of over-reporting was greatest for mathematics (about 1 semester) and science (about ½ semester).

C. Reliability.

Reliability coefficients for responses regarding familial background information, estimated from twin data, tend to be highly correlated with (r = .92) and slightly larger (.05, on the average) than validity coefficients estimated by comparing student and parent responses. Reliability coefficients are available for more variables than are validity coefficients because many student questionnaire items were not included in the parent questionnaire.

1. The coefficients for some of the major variables of interest that were not mentioned in preceding sections of the highlights are as follows:

—whether student was born in U.S.	1.00 (seniors)
—kindergarten attendance	.93 (sophomores)

*For a discussion of the concepts of validity, bias, reliability, and concistency, see section 2.5 on page 11."

—when began attending present school	.89 (seniors)
—student's race/ethnicity	.89 (seniors)
—number of siblings in high school next fall	.84 (seniors)
—school racial/ composition, grade 9	.83 (seniors)
—number of years student lived in U.S.	.81 (seniors)
—student's age	.78 (seniors)
—student's religious background	.78 (seniors)
—number of rooms in home	.71 (seniors)
—number of siblings in college next fall	.68 (seniors)

2. Students also were used as sources of information about their schools with regard to rule enforcement (five items), disciplinary problems (six items), and several other matters. The questions asked generally are of a highly judgmental nature (e.g., the extent to which "students talk[ing] back to teachers" is a problem). Consequently, most of the reliability coefficients estimated from twin data are only between .20 and .40.* While the coefficients for individual items for individual students are not very high, it should be noted that (1) many of these items can be combined into composite measures and (2) for many analytical purposes, it is appropriate to aggregate the data for many of these items (or composites formed from them) to the school level. Composites and school means, of course, have much higher reliability coefficients than those for individual items or individual students. (For example, if the reliability of a variable is .20 for an individual student, the reliability for the mean value of the variable for 36 randomly selected students in the school would be about .90.)

D. Consistency.

The extent of agreement between parent and child was examined for a number of variables for which the parental response could not be taken as a factual standard. Thus for these variables, the extent of agreement between parent and child reflects the degree to which they share the same perceptions and is *not* a measure of the quality of student responses. Of particular interest is the high correlation coefficient of .76 found for high school grade point average. The correlation coefficients for other variables, however, are lower; e.g., .61 for estimated schooling expenses for the coming school year, .07 to .47 for planned use of 13 possible sources for paying school and living expenses, .06 to .25 for estimates of costs of attending various kind of postsecondary education institutions, and .11 to .56 for the importance of various factors in choosing a college.

E. Quality of Data by Student Characteristics.

The quality of questionnaire responses was found to be better for some groups of students than for others:

1. Seniors generally provided better quality data than did sophomores. For example, the average validity coefficient for 12 family background items is .64 for sophomores, .67 for seniors; and the average reliability coefficient for 31 family background items is .66 for sophomores and .70 for seniors. With regard to school-related variables, the average reliability coefficients are .44 for sophomores and .54 for seniors.
2. Female students tended to provide slightly better quality data than males. The overall average validity coefficients are .64 for females and .61 for males for family background items, and .72 for females and .69 for males for high school grades and coursework.
3. White students provided better quality data than Hispanic or black students. For example, the mean validity coefficients for family background items are .61 for whites, .57 for Hispanics, and .54 for blacks; and for high school grades, the coefficients are .80, .66, and .65, respectively. Also, Hispanic and black students overstated to a greater degree than white students the amount of coursework they had taken in mathematics and science.
4. The quality of questionnaire data varies considerably with cognitive test performance level. From lowest to highest test score quartile, the mean validity coefficient rises from .56 to .67 for family background items and from .47 to .73 for amount of coursework taken. Some student questionnaire and transcript mean values also agree less closely for low- than for high-scoring students. The most extreme example is the amount of mathematics taken, where the mean calculated from questionnaires exceeds that calculated from transcripts by 1.8 semesters for students in the lowest test score quartile, but by only .6 semester for those in the highest quartile.

*The responses of twins, however, may agree more closely than would the responses of two randomly selected students in the same school because of the possible influence of attitudes and standards that are learned in the common home of twins.

Chapter 1 — Introduction

1.1 Background and Purpose

The base-year survey of the longitudinal study High School and Beyond (HS&B) took place in spring 1980. Data on a variety of topics were obtained by means of questionnaires and tests which were group-administered in school to national probability samples of about 30,000 high school sophomores and over 28,000 seniors. Although information is available about the reliability of the HS&B tests (Swineford, 1973; Heyns and Hilton, 1982), little information thus far has been available about the accuracy of the questionnaire data.[1]

HS&B provides a rare opportunity to examine the validity and reliability of questionnaire items completed by high school sophomores and seniors. This opportunity arises from three almost unique aspects of the study. First, in fall 1980 questionnaire data were collected from a subsample of over 6,500 parents of students who participated in HS&B. The major purpose of the parent survey was to garner information about how parents plan to finance the postsecondary education of their children. Another purpose, however, was to secure data that would enable NCES to assess the validity of student reports about family income, parental education, language spoken in the home, and other family and home variables. Second, in fall 1982 the high school transcripts of about 16,000 of the 1980 sophomore survey participants were obtained from their schools. The transcripts were collected primarily to secure information about course taking patterns and to relate these patterns to student and school characteristics and to test performance, college attendance, and other outcome measures. They also provide a check, however, on the accuracy of student reports of coursework taken and grades obtained. Finally, special provisions were made in HS&B to identify the co-twins of all sampled twins and to obtain questionnaire and test data from them. Through these efforts data were obtained from both members of about 500 sets of twins. The twin data are useful in providing estimates of the reliability of student reports about their schools as well as about their parents and home backgrounds.

Family background variables have been demonstrated to be important predictors of a child's educational attainment and success in school and of intergenerational occupational mobility (Averch *et al.*, 1971; Borus and Nestel, 1973). For this reason, many studies that employ HS&B data attempt to control for family background variables. Many analyses of HS&B data also rely heavily on student questionnaire data for values of various variables that reflect the student's school performance (e.g., grade average), educational experiences (e.g., kindergarten attendance and coursework taken), and school context (e.g., racial composition of classes in earlier grades, enforcement of rules, and discipline problems). Estimates of the reliability and validity of measures of family and school variables obtained from student questionnaires should enhance the researcher's ability to properly analyze HS&B data and to correctly interpret the results of the analysis.

Students were asked a variety of questions about their attitudes, plans, and expectations. It is not possible in this study to assess the quality of student response to questions of this nature. It is possible, however, to determine the extent to which parent and child see eye-to-eye regarding such matters as parental influence on student plans, the age at which the child will marry, the occupation the child will enter, the child's ability to complete college, attitudes about sex roles, college costs, and factors important in choosing a college. This report includes an analysis of the consistency of student and parent responses on these items. This kind of consistency information should provide insights regarding the extent of accurate communication of feelings, information, and intentions between generations within the family; and it should give indications of the degree of reality of student and parent perceptions of the world and of the child's future. Of special interest is the amount of consistency between parent and child perceptions of: (a) how much money the child will need

[1] Information about the reliability (internal consistency) of certain psychological and educational scales constructed from sets of questionnaire items utilized in HS&B as well as NLS-72, however, may be found in Riccobono *et al.*, 1981. The internal consistency of such scales is not examined in the present report.

for living and schooling expenses the year after high school; (b) the sources of funds (parents, student savings, government financial aid programs) for meeting these expenses; and (c) the importance in college selection of the cost and availability of financial aid at the college. Such data should be useful in postsecondary education financial aid studies.

1.2 Findings of Previous Studies

Several investigations of the quality of student questionnaire data obtained in the National Longitudinal Study of the High School Class of 1972 (NLS-72) have been conducted. The report of the results of one of these investigations (Conger *et al.*, 1976) is particularly valuable since it includes an extensive review and discussion of the literature on the validity and reliability of survey questionnaire research. Previous research involving NLS-72 and other data bases, however, generally has been more limited than that contained in the present report because of sample sizes that were too small to permit subgroup comparisons, restricted scope of variables studied, and problems of design and implementation.

1.2.1 Stability and validity study, 1972 seniors. Echternacht (1973) presented the results of "stability"[2] and validity studies conducted in conjunction with the NLS-72 base-year survey. Echternacht, based on a readministration of a subset of questionnaire items to a sample of 503 of the 1972 seniors, estimated reliability coefficients to be as follows (table 1.1):

Table 1.1. Estimated reliability coefficients by variable category, base-year NLS-72 student questionnaire.

Variable category	Range of coefficients
Persons influencing post-high school plans	.26 to .43
Attitudes toward self	.22 to .34
Attitudes toward high school	.19 to .30
Life values	.34 to .57
Future occupation	.57
Career selection values	.19 to .44
Life's work values	.12 to .35
Educational aspiration and plans	.42, .65
Race-ethnicity	.89

The NLS-72 base-year validity study compared the responses of 391 parents with those of their children (1972 seniors). Echternacht examined only the "bias" aspect of validity; that is to say, he analyzed the mean differences between parent and child responses rather than the correlations between the responses from the two sources. The bias in student responses was estimated for each category of about a dozen items. About 69 percent of the bias estimates were less than two percentage points. Some of the largest biases, in terms of the statistic he employed, were as follows (table 1.2):

Table 1.2. Item categories with large estimated biases, base-year NLS-72 student questionnaire

Item category	Difference (parent minus child) in pct. points
Length of residence in community = all his/her life	-7.7
English is the language most often spoken in home	5.1
Student has specific place to study	-6.7
Home has encyclopedia	-6.1
Family income is between $10,000 and $11,999	5.0
Family income is between $15,000 and $18,000	-10.4
Father's occupation was clerical	8.2
Father's occupation was operative	-5.4
Mother's highest education level = high school graduate	5.0

[2]This study was called a response stability rather than a reliability study because of the long lag (roughly 2 months) between the two data collection dates. The time lag made it questionable as to whether changes in responses were due to response unreliability or to actual changes in beliefs and attitudes.

The key assumption in a validity study of this kind is that parents' responses are better or more accurate than students' responses. Echternacht points out that his estimates of bias may have been adversely affected by the long lag (approximately 1 year) between student and parent data collections. He also indicates that the low parent response rate (74 percent) might have some biasing effect on the estimates.

1.2.2 Reliability of retrospective data. Lyons and Moore (1976) compared the responses given by a sample of 415 subjects in fall 1973 with their initial responses in spring 1972. The time referent for the questions asked in fall 1973 was spring 1972. The authors used the proportion of individuals who selected the same response in both surveys to measure reliability. This measure yields higher values than the measures of reliability conventionally employed.

The results of their study are summarized in table 1.3. For ordinal variables, the statistic shown in table 1.3 is the proportion of responses that agree to within one category.

Table 1.3. Proportion of responses in fall 1973 that agreed with initial responses in spring 1972: 1972 high school seniors

Variable (Time referent for all variables is spring 1972)	Agreement proportion
Father's, mother's educational attainment level*	.94, .95
Family yearly income*	.65
Items in home (11 items)	.67 - .98
Race-ethnicity	.95
Language background	.91
Religious background	.80
Limiting physical condition	.91
Type of community lived in*	.79
High school program	.72
High school grade average*	.93
Participation in extracurricular activities (9 items)	.77 - .93
Hours worked for pay*	.70
Father's, mother's educational desires for child*	.92, .94
When decided whether would go to college*	.82
Postsecondary plans of close friends	.64

*Response selections are ordinal. Statistic shown is the proportion of responses that agree to within one category.

1.2.3 Validity and reliability of survey research questionnaires. Conger *et al.* (1976) presented the results of a test-retest reliability study of a subset of 17 items from the NLS-72 second follow-up survey. Individuals who responded to the second follow-up survey by mail were mailed the short questionnaire containing the 17 items no earlier than 10 days after the completion date denoted on the second follow-up questionnaire. Reliability sample members who responded to the second follow-up survey by personal interview were reinterviewed with the short-form questionnaire 2 weeks later. The authors also reviewed the literature (33 references) on the quality of responses to NLS-type questions; and integrated their findings with those of other research. A validity study was not done; but previous research indicated to these authors that validity and reliability results generally are in close agreement. Based on the findings of other investigators, as well as on their own research, Conger *et al.* drew several major conclusions:

A. **Reliability and validity as a function of item characteristics.** The literature as well as their own study indicated that contemporaneous, objective, factually oriented items are more reliable and valid than subjective, temporally remote, or ambigious items. The NLS-72 second follow-up survey reliability study yielded estimated coefficients in the range .67 to .92 for contemporaneous, factually based items and between .36 and .86 for other items. Boruch and Craeger (1972) and van Es and Wilkening (1970) also found that items with a future as well as those with a retrospective orientation were less reliable than contemporaneous items. The literature indicates, however, that even factually oriented items differ in reliability depending on the importance (Astin, 1965), with personally important items being more reliable. Boruch and Craeger (1972) indicate that personally sensitive items (e.g., income) are less reliable than other factually oriented items. Finally, Conger *et al.* found that attitudinal and psychological variables were typically of moderate reliability.

The results of validity studies such as those by Walsh (1967 and 1968) and Kyaser and Summers (1973) were similar to those of reliability studies.

B. **Reliability and validity as a function of data collection procedures.** The results of the authors' reliability study showed that the interview procedure generally produced more reliable results than did the mail-in procedure. Studies by Walsh (1967, 1968) and Cannell and Fowler (1963) indicated that neither approach is consistently superior. Some important interactions were revealed. In particular, high ability and high SES persons were less influenced by data collection procedures than low ability or low SES persons. The latter groups were more cooperative and produced more accurate data in the interview procedure.

Previous researchers found a content by data collection procedures interaction. Questions which could be answered by consulting records were more accurately answered in a mail-in procedure.

C. **Reliability as a function of respondent characteristics.** Validity studies indicate that: (a) blacks provide less accurate information than whites (Borus and Nestel, 1971; Kerckhoff, Mason, and Poss, 1973; and Cohen and Crum, 1972); (b) low SES respondents furnish less accurate data than do high SES respondents (Borus and Nestel, 1971; Cohen and Orum, 1972; and Walsh and Burkhold, 1970); (c) students with higher grades produce more valid data than their counterparts (Boruch and Craeger, 1972; and Borus and Nestel, 1971); and (d) no clear edge exists for either males or females, but males tend to produce higher quality data on financial or numerical questions. The NLS-72 second follow-up survey reliability study results were in general agreement with these validity study findings with one major exception. The reliability study indicated a balanced set of racial differences with a tendency for minority youth to have higher reliability coefficients than whites.

1.3 Sources of Data

1.3.1 Twin component of student survey. The HS&B base-year survey employed a two-stage stratified sample. Public and private schools were stratified according to several key variables, and then schools within each stratum were selected with probabilities proportional to estimated average grade 10 and/or grade 12 enrollment. Certain types of schools were oversampled. In the second stage of the sample, simple random samples of 36 sophomores and 36 seniors, school size permitting, were chosen from each selected school. Detailed information about the basic HS&B sample design may be found in Frankel *et al.* (1981). Information about field procedures may be found in the student file user's manual (NORC, 1980); and some of the base-year survey results are described in a report by Peng *et al.* (1981).

Many researchers believe that twin data lend themselves to powerful analyses which enhance the utility of data from the overall study. Therefore, efforts were made to identify twins that had been among the cluster of 36 randomly chosen students. If the co-twin of a twin selected to participate in the survey attended the same high school and had not been already randomly selected into the sample, the co-twin was asked to take part in the survey. As a result of these efforts, questionnaires were completed by both members of the 276 sophomore and 235 senior twin sets.[3] The procedures used to identify twins and augment the sample by adding non-sampled co-twins are described more fully in the twin and sibling file user's manual (NCES, 1982).

1.3.2 Parent survey. A random sample of 312 of the schools that fully participated in HS&B was choosen for the parent survey.[4] In each parent survey school, simple random samples of 12 sophomores and 12 seniors were selected from those students who had completed HS&B questionnaires and taken the HS&B tests. A total of 3,654 sophomores and 3,547 seniors were selected by the procedures described in the parent questionnaire codebook (NORC, 1981).

The contractor attempted to collect data from the parents of these students by a combination of mailed-out questionnaires, telephone interviews, and personal interviews. All questionnaires were mailed to parents by October 10, 1980. Data collection ended December 31, 1980. A response rate of 91 percent was achieved; 3,367 forms were comleted by parents of sophomores and 3,197 by parents of seniors. Mothers filled in about 60 percent of the questionnaires, fathers about 35 percent, and other adults (stepparents, grandparents, etc.) the remaining 5 percent (table 1.4).

[3]These numbers exclude five sophomore and three senior twin sets where one student gave a negative response to the question, "Do you have a twin brother or a twin sister?"

[4]A school participated fully if it permitted tests as well as questionnaires to be administered.

Table 1.4. Relationship of parent questionnaire respondents to student: 1980 sophomores and seniors

Respondent	Sophomores		Seniors	
	Number	Percent	Number	Percent
Total	3,367	100.0	3,197	100.0
Mother	2,047	60.8	1,867	58.4
Stepmother	30	0.9	31	1.0
Father	1,157	34.4	1,152	36.0
Stepfather	38	1.1	32	1.0
Grandmother	26	0.8	30	0.9
Grandfather	7	0.2	4	0.1
Other relative or guardian	56	1.7	74	2.3
Unclassified	6	0.2	7	0.2

*Details may not sum to 100.0 due to rounding.

The composition of the realized parent survey sample by sex of student, student composite test score, family SES level (student-reported), and student racial/ethnic category is given in table 1.5. Note that the sample sizes are large enough for Hispanics and blacks to support the calculation of separate estimates of validity and consistency coefficients for each minority group for each cohort.

1.3.3 Transcript study. The sample for the transcript study consists of 18,152 selections from the 1980 sophomores who were eligible to participate in the first follow-up survey of HS&B.[5] Students who had attended private schools, high school dropouts, minority group students, and other segments of the population of special concern to education policymakers were heavily oversampled.

Table 1.5. Composition of parent survey realized sample by selected student characteristics: 1980 sophomores and seniors

Respondent characteristics	Sophomores		Seniors	
	Number	Percent	Number	Percent
Total	3,367	100.0*	3,197	100.0
Sex of student:				
Male	1,527	45.4	1,404	43.9
Female	1,649	49.0	1,690	52.9
Unclassified	191	5.7	103	3.2
Student test score quartile:				
Lowest	908	27.0	898	28.1
Second	825	24.5	780	24.4
Third	793	23.6	764	23.9
Highest	838	24.9	753	23.6
Unclassified	3	0.0	2	0.0
Socioeconomic status (student-reported):				
Lowest	881	26.2	972	30.4
Middle two	1,590	47.2	1,452	45.4
Highest	745	22.1	709	22.2
Unclassified	151	4.5	64	2.0
Race/ethnicity of student:				
Hispanic	350	10.4	298	9.3
Black	470	14.0	395	12.4
White	2,376	70.6	2,236	69.9
Other	113	3.4	143	4.5
Unclassified	58	1.7	125	3.9

*Details may not sum to 100.0 due to rounding.

[5]This sample was augmented by the addition of 275 non-sampled co-twins.

Transcripts were provided by high schools in fall 1982 for 15,941 (88 percent) of the 18,152 individuals for whom they were requested. Courses were assigned a six-digit classification code based on a taxonomy developed by Evaluation Technologies Incorporated for NCES,[6] and course grades were coded by NORC so as to be on a uniform scale regardless of the school's grading system.[7]

The investigation of the quality of student questionnaire responses utilizes transcripts only for those students who were still attending high school in spring 1982. The composition of this subset of the realized sample of transcripts is shown in table 1.6. The sample sizes are large enough to obtain very accurate estimates of the quality of data provided by subgroups of students, including Hispanics and blacks.

Table 1.6. Composition of transcript study realized sample by selected student characteristics: 1980 sophomores attending high school in 1982

Student characteristics	Number	Percent
Total	13,173	100.0
Sex:		
Male	6,545	49.7
Female	6,628	50.3
Unclassified	0	0.0
Test score quartile (1980):		
Lowest	2,324	17.6
Second	2,695	20.5
Third	2,890	21.9
Highest	3,434	26.1
Unclassified	1,830	13.9
Socioeconomic status (SES) quartile (1980):		
Lowest	2,829	21.5
Middle two	5,615	42.6
Highest	3,246	24.6
Unclassified	1,483	11.3
Race/ethnicity:		
White	8,049	51.1
Hispanics	2,692	20.4
Black	1,779	13.5
Others	536	4.1
Unclassified	117	0.9

[6]Ludwig *et al.*, 1982.

[7]Jones *et al.*, 1983.

Chapter 2 — Methodology

2.1 Variables Analyzed

Table 2.1 lists the variables analyzed in the parent-child portion of this study. The variables are grouped into two broad categories. The first contains variables that are factual in nature and for which the parent's response provides a standard for judging the accuracy of the student's response. The second category includes all other variables. For simplicity, the first category is labeled "validity (parent-referent)," the second "consistency."[8] The distinction between these categories may be blurred for some variables. It is not always entirely clear whether the parental response provides a sound factual standard, especially where an element of judgement is involved. While of conceptual importance, the distinction has little bearing on how the data are analyzed. Both validity and consistency are measured in the same way—by the degree of correlation between parent and child responses.

Table 2.2 lists the variables that were included in the reliability (twin/co-twin) portion of this study. Many more variables could be analyzed in the reliability portion of the study than in the validity and consistency components. The variables were divided into two categories, home/family and school.

Transcript data were employed to determine the validity of student responses regarding high school grade average, number of semesters of coursework taken in each of seven areas, and whether each of six courses had ever been taken.

2.2 Treatment of Data

Cases were omitted from the analysis whenever data were missing due to instrument or item nonresponse or were reported in an invalid way for one or both of the two sources being compared. Cases also were omitted from the analysis whenever at least one member of a child-parent or twin/co-twin pair answered "don't know" or "not applicable" to a variable otherwise measured on an ordinal scale. For the question concerning the age at which the student expects to (or is expected to) get married, have first child, etc., cases with "don't expect to do this" responses by the child, parent or both also were omitted from the analysis.[9]

Both members of the twin pairs answered exactly the same questions. The format and wording of some of the items in the parent questionnaire, however, differed from that for corresponding items answered by students. The general strategy employed where this occurred was to try to map or transform parent responses into the student questionnaire format since it is the quality of student responses that is being assessed. The two most pronounced instances of this nature are the parental occupations and family income questions. Parents wrote in the names of occupations, which were assigned three-digit Census occupational classification codes by the National Opinion Research Center (NORC). The three-digit codes then were matched by NORC to the 16 occupational categories listed in the questionnaire answered by students. Income information was obtained from parents by having them write in code letters corresponding to 19 class intervals ranging from none to $500,000 or more. The midpoints of the designated class intervals were summed for four items: respondent's wage/salary income, respondent's business/farm income, spouse's wage/salary income, and spouse's business/farm income. The sums then were mapped into the seven family income class intervals given in the questionnaire answered by students.

[8]Actually, however, for one of the variables in the consistency category (namely, mother's aspirations for child's education) the mother's response can be considered to provide a standard for assessing the quality of the student's response.

[9]The percentage of parents who gave this answer was 2 percent or less for each sub-item. For students, the percentages who gave this answer ranged from 2 percent for "first job" and "live in own home" to 10 percent for "have first child."

Table 2.1. Questionnaire item numbers of variables analyzed in study of validity of student responses and consistency of parent-child responses: 1980 sophomores and seniors* **

Variable	Questionnaire item number			
	Sophomores		Seniors	
	Student	Parent	Student	Parent
A. Validity (parent referent)				
Items in home (8 items)	103	59	N/A	N/A
Mother's education	42	29/44	42	38/50
Father's education	39	29/44	39	38/50
Mother's occupation	41	38/46	41	47/53
Father's occupation	38	38/46	38	47/53
Family income	100	65/66	101	70/71
Own/rent home	101	60	102	64
Home language (2 items)†	14/15	57/58	14/15	62/63
Number of siblings (2 items)	N/A	N/A	98/99	60
Mother worked during various periods (3 items)	37	50	37	55
Race/ethnicity	90/91	56	89/90	61
B. Consistency				
High school grade average	7	3	7	3
Mother's influence on plans	N/A	N/A	49	16/54
Father's influence on plans	N/A	N/A	49	16/54
Age student expects to . . . (5 items)	78	14	81	19
Student's job plans	68	8	62	8
Student's college plans (4 items)	72	9	68	9
Mother's aspirations for child's education	70	5	66	5
Ability to complete college	73	10	69	10
Sex role attitudes (3 items)	63	41	N/A	N/A
Estimated living expenses	N/A	N/A	76	12
Source of money for expenses (13 items)	N/A	N/A	79	15
Estimated cost of postsecondary education (3 items)	111	20	111	24
College choice factors (7 items)	N/A	N/A	116	37

"*N/A *Indicates that the item did not appear in both student and parent questionnaires.

**The questionnaires employed in the surveys are not reproduced in this report, but copies of them are available upon request from NCES.

†Student item numbers refer to the Student Identification Pages (language file).

In order to assess the quality of student questionnaire reports regarding high school grade averages and coursework, corresponding statistics had to be calculated for each student from the detailed information in the transcript file. Information about how this was done is given in section 3.2 of this report.

2.3 Sample Sizes

The numbers of observations upon which the estimates of child-parent correlation coefficients were based are given in tables A-1, A-7, and A-9. For most variables, the numbers are quite high, ranging up to 95 percent of the maximum possible (viz., 3,367 for sophomores and 3,197 for seniors). Table 2.3 lists the variables for which the number of cases used in the child-parent calculations fell below 80 percent of the maximum. When

data for a child-parent pair were not usable, it generally was because the student gave a response such as "don't know" or did not answer the question at all.[10,11]

Table 2.2. Questionnaire items included in reliability study: 1980 sophomores and seniors*

Variable	Questionnaire item no.	
	Sophomores	Seniors
Home/family variables:		
Father/mother lives in same household (2 items)	36	36
Mother worked (3 items)	37	37
Father's job category	38	38
Father's education	39	39
Years father has lived in U.S.	40	40
Mother's job category	41	41
Mother's education	42	42
Years mother has lived in U.S.	43	43
Whether student was born in U.S.	44	44
Years student has lived in U.S.	45	45
Student's age	85	84
Racial/ethnic group	90/91	89/90
Race	90	89
Religious background	92	91
Family income (one-thirds)	99	100
Family income (one-sevenths)	100	101
Family owns/rent house or apartment	101	102
Number of rooms in home	102	103
Items in home (9 items)	103	104
Number of siblings in college next fall	–	98
Number of siblings in high school next fall	–	99
School variables:		
Number of school changes due to family moving	11	–
When began attending this school	–	10
Whether attended kingergarten	12	–
Instruction methods used in school courses (6 items)	–	3
School disciplinary problems (6 items)	19	–
Enforcement of school rules (5 items)	20	–
School ratings (8 items)	57	53
Whether school has minimum competency tests	58	54
School factors that interferred with education (5 items)	–	52
Number of black students in classes (3 items)	104	105
Number of Hispanic students in classes (3 items)	105	106

*The questionnaires employed in the surveys are not reproduced in this report, but copies of them are available upon request from NCES.

The subset of students who got to an item and felt they had enough knowledge to answer the question in a specific way, probably provided better quality data than the other students. Hence, the estimated validity and

[10] A major reason why students did not answer certain questions is that some students did not have sufficient time to complete the questionnaire. The sophomore (senior) questionnaire had 114 (115) items to be answered by all students. About 96 percent (97 percent) of the sophomores (seniors) got as far as item 80, but only about 87 (89 percent) made it all the way to item 114 (115).

[11] In three instances which are not listed in table 2.3, the number of observations utilized in the computations was constrained by other considerations: (a) the analysis of the items regarding estimated cost of schooling and college choice factors was limited to those cases where the seniors indicated possible college attendance and the parents reported school attendance by the child—about 2,000 usable cases; (b) the analysis of the items regarding the mother's aspirations for her child's education was restricted to cases where the mother was the parent questionnaire respondent - about 1,600 usable cases for the sophomore cohort, 1,550 for the senior cohort; (c) the analysis of the language background items was limited to instances where the students provided some evidence of a foreign language background—about 600 usable cases for each cohort; and (d) the analysis of items about whether the mother had a job at various times in the past was restricted to those cases (about 2,100 sophomores and 1,900 seniors) where a female answered the parent questionnaire.

consistency coefficients for variables with high percentages of unusable cases may present a somewhat rosier picture than is warranted.

Table 2.3. Proportion of total cases used in child-parent analysis for variables for which the proportion was less than .80 of the maximum number of cases for either cohort

Variable	Cohort	
	Sophomore	Seniors
Mother's education	.78	.85
Father's education	.63	.70
Mother's occupation	.57	.62
Father's occupation	.66	.68
Family income	.79	.83
Father's influence on plans	-	.74
Age expect to . . . (5 items)	.61 to .82	.56 to .82
College expectations in earlier grades (3 items)	.65 to .78	.69 to .75
Sources of money for living expenses (13 items)	-	.60 to .76

The proportion of the 276 (235) sophomore (senior) twin file cases utilized in the reliability portion of this study exceds .80 for 63 of the 72 items analyzed (table 2.2). The 9 exceptions are identified in table 2.4.

Missing data is not a problem for high school grades and coursework. All validity estimates made in the portion of the study employing transcripts are based on at least 90 percent of the maximum of 13,173 cases.

Table 2.4. Proportion of total cases used in twin analysis for variables for which the proportion was less than .80 of the maximum number of cases for either cohort

Variable	Cohort	
	Sophomore	Seniors
Father's education	.51	.68
Mother's education	.64	.80
Family income (one-thirds)	.69	.77
Family income (one-sevenths)	.72	.80
Mother worked before elementary school	.73	.77
Hispanic students in grades 1, 6, and 9	.76 to .77	.82 to .86
Electric dishwasher	.78	.86

2.4 Subgroup Comparisons

When possible, separate estimates of reliability, validity, and consistency coefficients were calculated for each cohort.[12] Older students should be better informed than younger ones about such matters as parental education and family income; and it was anticipated that high school seniors would have firmed-up their plans and attitudes and be more realistic about them. Hence, the degree of consistency between data for students and their parents should be greater for seniors than for sophomores.

For each cohort, validity and consistency coefficients were estimated for subgroups of students. These subgroups were formed by classifying students by sex, race/ethnicity (student's report), achievement level (measured by their composite scores on HS&B verbal and quantitative tests), and parental socioeconomic status (SES) level (student's report). Based on previous research findings, it was hypothesized that for most variables, data quality would be higher for white than for minority-group students, higher for high- than for low-achieving students, and higher for high- than for low-SES students. In addition, separate estimates of validity and consistency coefficients were made on the basis of whether the father (or another male) or the mother (or another female) completed the parent survey form. It was expected that for some variables, such as mother's education level and sex role attitudes, there would be closer agreement between mother-child

[12]The transcript study was limited to the 1980 sophomore cohort.

than father-child responses; while for family income, father's occupation, and some other variables, it was thought that there would be better father-child than mother-child agreement.

5 Measurement of Validity, Bias, Reliability and Consistency

Psychologists commonly use the term validity coefficient to refer to the correlation between the observed variable and the construct or true score (Werts, 1974). Although information provided by parents about home and family background variables is not infallible, parent responses are assumed to be much more accurate than those of their children; and parent responses on such matters are the standard against which the validity of student responses are judged in this report. The correlation coefficient between parent and child responses, however, undoubtedly somewhat understates the validity of student responses. This is due to the fallibility of parent responses and the lag of about 6 months between the collection of data from parents and students.

Fallibility of the standard would seem to be less of a problem when judging the quality of student responses against high school transcript data. Nevertheless, there is a certain amount of judgment involved in classifying and coding some courses; and inter-coder reliability is not perfect. Moreover, in mapping transcript data into the course categories in the student questionnaire, there is some ambiguity. For example, if the student is asked whether he or she has ever taken trigonometry, it is not entirely clear how the accuracy of the student's answer should be judged if the student took a unified mathematics course that included a trigonometry component but did not take a separate course in trigonometry.[13] For reasons such as these, this study's estimates of the validity of student reports about grades and coursework also should be considered as lower bounds on the actual validity of student responses.

The concept of "bias" also plays a role in the assessment of the accuracy of student responses. Bias is measured by the difference in means between student responses and those obtained from the standard (parent or transcript) while validity is measured by the correlation in responses from the two sources. Thus the total error of student reports of parental income, for example, may be thought of as the root mean square of the sums of two components: the variable error as measured by the correlation coefficient and the square of the systematic error or bias in student reports (Kish, 1965).

Reliability is determined by the consistency of repeated independent measurement of a fixed value with the same instrument. A traditional test theorist considers the correlation between equivalent, parallel, or alternate forms of a test to be the reliability coefficient (Werts, 1974). Most investigations of the reliability of questionnaire responses have used the "test-retest" approach (Conger, 1976). The expected value of the response is assumed to be the same at both repetitions. The farther apart the two repetitions are spaced in time, however, the more likely that a difference between the two responses reflects true change rather than unreliability. On the other hand, the closer together the two administrations of the instrument, the more likely the respondent is to recall and therefore repeat the previous response. The former situation results in an underestimation of response reliability, the latter results in an over-estimation.

The correlation between responses of twins who were administered questionnaires at the same session should be a better estimate of the reliability of certain items than the estimate resulting from the test-retest approach. The test-retest approach, on the other hand, provides reliability estimates for attitudinal variables while the approach utilizing twin data does not. The twin data are employed in this study to estimate the reliability coefficients for home and family background variables and certain present and past school experiences. All twin pairs were attending the same school and living in the same household in spring 1980. The estimates of reliability coefficients assume, for some variables, that both members of the pair shared common school experiences (e.g., kindergarten attendance) and home experiences at earlier stages of their lives as well. Estimated coefficients also are presented for a few school variables (e.g., instructional methods used in courses taken) for which the true value might not always be the same for both members of the twin pair. For these variables, the correlation between twin respones sets a lower bound for the reliability coefficient.

Finally, we have taken the opportunity to examine the consistency or congruence between parent and child on certain attitudinal variables (e.g., factors important in choosing a college). The extent of agreement between parent and child reflects neither the validity nor the reliability of student responses. It simply indicates the extent to which parent and child see eye-to-eye on such matters.

---

[13] this study, such students were judged to have taken trigonometry.

The student questionnaire employed only closed-ended items. Respondents were not asked to write in numerical values or other kinds of answers, but to mark one of several mutually exclusive, exhaustive response categories. Imprecise, but rank-orderable, quantifying words frequently were used to define the categories (e.g., agree strongly, agree, disagree, and disagree strongly). In instances where response categories were defined quantitatively (e.g., family income categories), at least one category had an open-ended interval (e.g., family income of $38,000 or more). Thus all measurements were on a nominal or ordinal scale, even though the underlying variable often was continuous.

It should be noted that our analysis examined the overall relationships of the patterns of responses of the student and the standard, and did not attempt to examine the extent of association at the extremes or at various other points of the distributions. For example, it presents the correlation between parent and child for the entire range of educational attainment levels of the father. The full distribution is not decomposed with a separate measure of association calculated for various dichotomies, such as advanced degree vs. all other levels. The complete percentage distributions are tabulated, however, and are discussed in the sections of the report that deal with the topic of bias. The discussion of bias, of course, concentrates mainly on differences between the means of the distributions for the students and the parents (or transcripts).

The same measures of association were used for measuring reliability and the correlational aspect of validity and consistency. For variables measured on a fully-ordered scale (e.g., family income and number of siblings), Pearson's product-moment correlation coefficient (r) was employed. For variables measured on a nominal scale (e.g., race/ethnicity and religious background), the statistic called Cramer's V was employed (Kendall & Stuart, 1979; Bishop *et al.*, 1975). Like r, V can reach a maximum value of 1; and for dichotomous variables, V equals r. Bias was measured by the difference in means or proportions.

Chapter 3 — Findings

This chapter is organized into three major sections based on classification of variables into three categories: family background variables, school variables, and all other variables.

In section 3.1, student questionnaire responses to family background variables are examined carefully—first, with regard to validity and bias using parent data as the standard and then, with regard to reliability using the twin file. Next, the relation between reliability and validity coefficients is investigated for variables for which both kinds of coefficients were obtained. The quality of family background data provided by seniors then is contrasted with that furnished by sophomores; and the validity of questionnaire responses is compared for subgroups of students which were formed by classifying cases by whether a male or female answered the parent questionnaire and by the student's sex, race/ethnicity, cognitive test performance level, and socioeconomic status (SES). Detailed tabulations of validity, bias, and reliability coefficients for family background items may be found in appendix tables A.1, A.2, and A.3, respectively.

The investigation in section 3.2 of the quality of data about school-related variables has two major components. The first involves estimating the accuracy of senior reports of high school grades and coursework by comparing student reports with values derived from their transcripts. The second involves estimating the reliability of a variety of variables by comparing responses of twins. The transcript sample size is sufficiently large to provide accurate estimates of data quality for various subgroups of students, but the twin sample is not. Validity coefficients are presented in table A.4, bias values in table A.5, and reliability coefficients in table A.6.

Finally, in section 3.3 the consistency of parent and student responses to a variety of variables is investigated, as is the mean difference in response from the two sources. For these variables, with one exception, the parent's response cannot be taken as a standard for judging the accuracy of the child's response. The analysis in this section merely indicates the extent to which the parent's views conform with those of the child. Comparisons of parent-child consistency coefficients are made for the same subgroups employed in section 3.1. Consistency coefficients and mean difference values for grade-point average and variables involving attitudes and expectations are given in tables A.7 and A.8; similar coefficients for variables involving financial and college matters are presented in tables A.9 and A.10.

Table A.11 contains approximate values of standard errors for the estimated correlation coefficients presented in prior appendix A tables and in the tables in this chapter. Tables A.12 and A.13 give approximate standard errors for bias estimates.

3.1 Family Background Variables

3.1.1 Validity—parent as standard. Estimated validity coefficients for student responses to family background questionnaire items cover a wide range of values, roughly from .20 to .90 (table 3.1).

The validity coefficients for student reports of the educational attainment levels of their parents are quite high (.80 to .90). These coefficients, of course, apply only to the subset of students who felt they knew the answers to the questions.[14]

The validity coefficients for father's occupation (.57 and .61), mother's occupation (.44 and .45), and family income (.50 and .59) are of modest size. The family income questions asked students, however, differed considerably from those asked parents (section 2.2); and the results of the reliability study, which are presented later (section 3.1.3), would seem to indicate that, as a consequence, family income validity coefficients may have been substantially underestimated here.

All 4 items with coefficients less than .40 and 7 of the 15 with values less than .70, are "household possessions" items. These items were not intended to be used individually, but to be employed, along with the "typewriter" and "electric dishwasher" items, in a household possessions composite. The coefficient for this composite was not calculated, but it should be higher than the average value of the component items.

[14] In the base-year survey of HS&B, 18 percent of the sophomores and 11 percent of the seniors responded "don't know" to the question about their father's education. For mother's education, the number of "don't know" responses was almost as high, 17 percent for sophomores and 10 percent for seniors.

Table 3.1. Estimated validity coefficients for 1980 sophomore and senior responses to family background questionnaire items

Variable	Validity coefficient	
	Sophomores	Seniors
High coefficients (.71 to .90):		
Father's, mother's education level	.87, .81	.89, .85
Electric dishwasher in home	.85	–
Siblings in high school next fall	–	.85
Socioeconomic status (SES) composite	.80	–
Race/ethnicity*	.78	.79
Language usually spoken in home**	.74	.78
Mother had job when child was in high school†	.70	.72
Moderate coefficients (.50 to .70):		
Typewriter in home	.63	–
Mother had job when child was in elementary school†	.62	.65
Two or more cars or trucks that run	.57	
Own or rent house or apartment	.56	.58
Daily newspaper in home	.54	–
Mother had job before child was in elementary school†	.53	.53
Other language usually spoken in home**	.50	.61
Family yearly income	.50	.59
Low coefficients (.20 to .49):		
Mother's occupation	.44	.45
Pocket calculator in home	.39	–
More than 50 books in home	.35	–
Encyclopedia, other reference books in home	.35	–
Specific place to study in home	.21	–

*This is not truly a validity coefficient since parents were asked to identify their own race/ethnicity, not that of their child.

**Analysis includes only students who indicated they had some foreign language background.

†Analysis includes only those cases where a female answered the parent questionnaire.

The socioeconomic status (SES) composite is an equally weighted average of the household possessions composite and four other variables: father's occupation, father's education, mother's education, and family income. Since the composite was computed for every student who provided data for at least two of the five components, SES composite values are available for about 95 percent of all base-year HS&B participants. This composite has been widely used since it sums up a number of aspects of the student's background in one statistic and has a value (i.e., is not missing) for nearly all cases. Thus, it is encouraging that the validity coefficient for the composite is .80 for sophomores and undoubtedly somewhat higher (probably about .84) for seniors.[15]

Validity coefficients for seniors for variables that do not enter into SES composite calculations are as follows: (a) .85 for number of siblings in high school next fall;[16] (b) .79 for race/ethnicity; (c) .78 for language usually spoken in home and .60 for other language spoken in home; (d) .58 for whether home or apartment is owned or rented; and (e) about .72, .65, and .53 for whether the mother worked full-time, part-time, or not at all during three time periods ranging from high school to before elementary school.[17]

The results of subgroup comparisons are summarized in table 3.2 and are further illustrated in figure 1. Figure 1A contrasts the coefficients for seniors with those for sophomores. A point above the line indicates that the coefficient for seniors is larger than the one for sophomores for the same variable. Figures 1B-1F

[15] As will be seen in the following paragraph and subsequent sections, seniors consistently provided more valid and reliable data than sophomores.

[16] The coefficient for number of siblings in college is only .60; but this is an underestimate because the corresponding parent questionnaire item refers to all children, not just to siblings of the HS&B student, and includes all postsecondary institutions, not just colleges.

[17] See footnotes to table 3.1.

*Errata

5. Page 14, Table 3.1, middle section. Insert the two lines

"Father's occupation	.57	.61"
"Siblings in college next fall	–	.60"

contrast the validity coefficients for subgroups of students defined by the student's sex, race, test performance, and SES and by the parent questionnaire respondent's sex. Data for both sophomores and seniors, identified by different symbols, are plotted in figures 1B-1F.

In most instances, but not all, subgroups were found to differ in the ways hypothesized (section 2.4).

Cohort. As expected, seniors provided more valid family background data than did sophomores.[18] It can be seen from figure 1A that all points are above or on the straight line. Thus, for all 12 variables for which data are available for both cohorts, the estimated coefficient for seniors is at least as high as the corresponding one for sophomores. The average coefficient is .635 for sophomores and .671 for seniors.

Table 3.2. Mean validity coefficients for family background variables for subgroups of 1980 sophomores and seniors, by selected student and parent characteristics

Subgroup	No. of coefficients averaged	Mean validity coefficient
Cohort:		
Sophomore	12	.635
Senior	12	.671
Sex of students:		
Male	35	.613
Female	35	.639
Race-ethnicity of students:		
White	29	.612
Hispanic	29	.570
Black	29	.544
Cognitive test performance:		
Lowest quartile	35	.562
Highest quartile	35	.669
Socioeconomic status:		
Lowest quartile	34	.547
Highest quartile	34	.588
Sex of parent questionnaire respondents:		
Female	29	.612
Male	29	.635

Sex. Perhaps surprisingly,[19] female students furnished more valid responses to questionnaire items concerning family background than did male students (figure 1B). The difference in mean values is small (.639-.613 = .026), but highly significant from a statistical point of view (table 3.3).

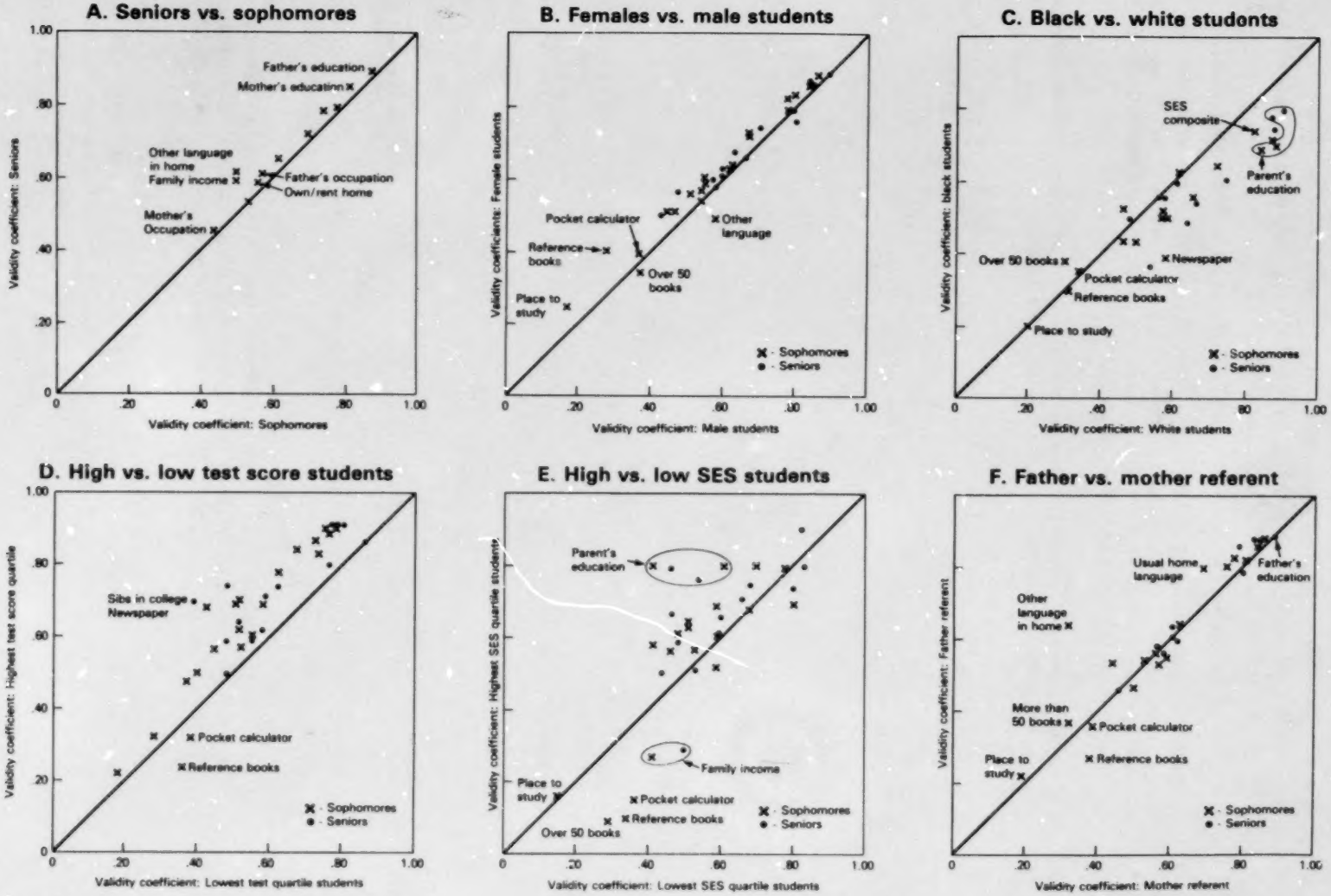
Academic ability. Students who performed well academically, as measured by scores on the HS&B cognitive tests, answered questions about family background matters in a much more valid way than did students who performed poorly (figure 1D). The difference in mean coefficients is substantial (.669- .562 = .107).

Socioeconomic status (SES). The results by SES composite (figure 1E) are mixed. For many variables, especially parental education level, the validity coefficients are higher for high- than for low-SES students; but for other variables, most notably family income, the coefficients are higher for low- than for high-SES students. Because many of the family background variables enter directly into the calculation of the SES composite, however, it is difficult to attach much meaning to comparisons based on SES level.

[18] See table 3.3 for results of statistical tests of significance.

[19] Boruch and Creager (1973) found no male male difference in validity and Cohen and Orum (1972) found that neither sex produced consistently superior data.

Figure 1.—Estimated validity coefficients for student responses to family background questionnaire items: 1980 sophomores and seniors*



*See tables 3.3 and A-1.

BEST COPY AVAILABLE

Table 3.3. Results of statistical tests of significance comparing subgroups with regard to validity, reliability, and consistency of questionnaire data and comparing validity with reliability coefficients: 1980 sophomores and seniors.

Figure number and comparison group (Groups with significantly higher coefficients in italic)	Sample size	Sign test a/			
		Number of coefficients relative to zero-difference lines (see figures)			Statistical significance b/
		Above line	Below line	On line c/	
Validity—Family background items					
1A. *Seniors* vs. sophomores	12	11	0	1	**
1B. *Female* vs. male students	35	27	3	5	**
1C. Black vs. *white* students	29	5	22	2	**
1D. *High* vs. low test score	35	32.	2	1	**
1E. *High* vs. low SES students	34	23	10	1	*
1F. *Father-child* vs. mother-child	29	17	8	4	†
Reliability—Family background items					
2. *Seniors* vs. sophomores	31	22	8	1	**
Family background items					
3. *Reliability* vs. validity	31	25	4	2	**
Validity—School-related items					
4A. *Female* vs. male students	14	12	1	1	**
4B. Black vs. *white* students	14	0	14	0	**
4C. *High* vs. low test score	14	13	1	0	**
4D. *High* vs. low SES students	14	10	3	1	*
Reliability—School-related items					
5. *Seniors* vs. sophomores	15	14	1	0	**
Consistency—Grades, attitudes, plans					
6A. *Seniors* vs. sophomores	10	9	1	0	*
6B. Female vs. male students	28	13	12	3	NS
6C. Black vs. *white* students	28	3	25	0	**
6D. *High* vs. low test score	28	18	9	1	†
6E. High vs. *low SES* students	28	9	18	1	†
6F. Father-child vs. mother-child	26	12	12	2	NS
Consistency—Financial and college items					
7A. *Female* vs. male students	28	19	8	1	†
7B. Black vs. *white* students	28	5	23	0	**
7C. *High* vs. low test score	28	22	5	1	**
7D. *High* vs. low SES students	28	20	8	0	**
7E. *Father-child* vs. mother-child	28	21	6	1	**

a/As applied here, the sign test is a non-parametric test based on the probability associated with the occurrence of a particular number of points above (plus) and below (minus) the zero-difference line as determined by reference to the binomial distribution (Siegal, 1956).

b/One-tailed tests of significance were employed in all comparisons except those based on the sex of the student and sex of the parent for grades, attitudes, and plans. The symbols †, *, and ** indicate statistical significance in the direction hypothesized, if any, at the .10, .05, and .01 levels, respectively; and NS means not significant at the .10 level. It was hypothesized that seniors and white, high test score, and high SES students provide better quality and more consistent data than sophomores or black, low test score, or low SES students. It also was hypothesized that there would be better agreement between student responses and those of their fathers than those of their mothers for family background and financial and college items. Furthermore, it was hypothesized that reliability coefficients would be higher than validity coefficients.

c/A point is considered to be "on the line" if its coordinate values, rounded to two decimal places, are equal.

Parent questionnaire respondent's sex. Thus far the discussion has been about the validity of student questionnaire responses as a function of student characteristics. The parent's response has been the standard against which the validity of the child's response has been judged. Parent responses, of course, actually are of varying degrees of accuracy. It was thought that the father might provide more accurate information than the mother for many family background items. Thus, validity coefficients for those cases where mothers (or other females) had answered the parent questionnaire were compared with those cases where the respondent had been a male (figure 1F).[20] It was found that there was a tendency for male-referent coefficients to be slightly higher than female-referent ones, the two sets of coefficients averaging .635 and .612 respectively. This difference also can be considered as evidence of a fallible standard, and thus to be indicative that the coefficients calculated in this study understate the actual quality of student questionnaire data.

3.1.2 Bias—parent as standard. The previous section examined the degree of correlation between the questionnaire responses of children and their parents. This section looks at differences in response means of children and parents. If the mean for students is higher than that for their parents, this is evidence of a positive bias. A positive bias is associated with over-reporting by the student, a negative bias with under-reporting.

Table A-2 shows that sophomores under-reported that they have a specific place to study (48 vs. 57 percent) and that the home has a pocket calculator (80 vs. 87 percent), but over-reported that the family has two or more cars or trucks that run (79 vs. 75 percent) and that the home has 50 or more books (85 vs. 81 percent).[21]

Students over-reported that their parents have a high school education only (45 vs. 36 percent for sophomore mothers), but under-reported that their parents have had some but less than 2 years of postsecondary education (10 vs. 24 percent for sophomore parents). The size of the bias in student reports of parental education tends to be somewhat smaller for father's than for mother's education level, and smaller for seniors than for sophomores.

Students tended to underclassify their father's occupation as clerical (2 vs. 5 percent, sophomores) and craftsman (17 vs. 23 percent, sophomores), but to overclassify their fathers into the categories of farmer (5 vs. 2 percent) and laborer (10 vs. 5 percent). Children underclassified the occupations of their mothers as clerical (31 vs. 35 percent, sophomores), but overclassified their mother's occupation as laborer (3 vs. 1 percent) and professional or school teacher (23 vs. 17 percent).

Students under-reported that the family income was either very low or very high, instead over-reporting it to fall within an intermediate range. For example, 68 percent of sophomores but only 40 percent of their parents reported income values between $7,000 and $25,000 per year. Overall, children underestimated the income of their family by an average of about $3,000 or 12 percent.[22]

Those students who indicated they come from a foreign-language background tended to over-report that the language is Spanish (31 vs. 27 percent, sophomores).

These students, when asked whether another language also was spoken in the home, tended to over-report that there was a second language (78 vs. 72 percent, sophomores) and to identify it as Italian, French, or German (18 vs. 10 percent, sophomores). In most cases, the biases were smaller for seniors than for sophomores.

3.1.3 Reliability—twin file. Table A.3 presents estimated reliability coefficients for 33 family background variables, a much larger number than could be included in the validity study.[23] Thus, use of twin data allows a more extensive investigation of data quality. Forty-two percent of the coefficients are at least .75 in value, 44 percent are between .50 and .74, and only 14 percent are less than .50. Reliability coefficients are summarized in table 3.4. Note that for seniors the coefficients are about .90 for parental education level, .75 for family income (seven categories), .65 for father's occupation, and .86 for the SES composite.

---

[20]There is evidence that in some instances, while the mother was the designated respondent, the father actually completed the section of the questionnaire dealing with financial matters. No attempt was made in this study to adjust for those circumstances.

[21]All of the bias estimates pointed out in this section of the report differ from zero by at least two standard errors. (See table A.12 for standard error values).

[22]If other income items such as dividends, interest, and rent had been included in the estimate of family income, the overall bias would have been smaller. Dresch *et al.* using the broader definitions, found a mean of $28,000, which was quite close to the U.S. Bureau of Census 1979 Current Population Survey estimated mean income for families headed by persons age 45 to 54 of $28,200.

[23]Some of the variables in the reliability analysis could not be included in the validity analysis because the information was not asked for in the parent questionnaire.

Table 3.4. Summary of estimated reliability coefficients for 1980 sophomore and senior responses to family background questionnaire items (twin file)

Variable	Reliability coefficient	
	Sophomores	Seniors
Socioeconomic status composite	.82	.86
Father's, mother's education level	.94, .85	.89, .90
Father's, mother's occupation	.56, .51	.65, .53
Family income (three, seven categories)	.66, .69	.71, .75
Electric dishwasher in home	.85	.81
Two or more cars or trucks that run	.72	.70
Typewriter in home	.70	.75
Other items in home	.19 to .57	.30 to .60
Number of rooms in home	.64	.71
Own or rent house or apartment	.58	.60
Father, mother lives in household	.84, .44	.86, .51
Mother worked at various points in time	.62 to .78	.63 to .84
No. of years parents, student lived in U.S.	.70 to .79	.77 to .81
Student was born in U.S.	.96	1.00
Student's age, race, religion	.69 to .82	.78 to .89
No. of siblings in college, high school next fall	-	.68, .84

As anticipated, the questionnaire responses of seniors tended to be more reliable than those of sophomores. (See figure 2.) The means of the reliability coefficients for the 31 variables measured for both cohorts are .701 for seniors and .664 for sophomores. Earlier (section 3.1.1) it was pointed out that seniors provided more valid responses than did sophomores.

3.1.4 Comparison of validity and reliability coefficients. Both reliability and validity coefficients were estimated for 19 sophomore cohort and 12 senior cohort family background variables. Conceptually, reliability sets an upward bound for validity (Werts *et al.*, 1974). This theoretical relationship between reliability and validity is supported by our data (figure 3).[24] The two statistics are highly correlated (r = .915), and the mean reliability coefficient is higher than the mean validity coefficient (.671 vs. .619).[25]

Thus the validity of student questionnaire responses for many variables other than those covered in the parent questionnaire can be approximated from the corresponding reliability coefficients. This should hold true for school-related as well as for family background variables.

3.2 School-related Variables

3.2.1 Validity—transcript as standard. The quality of senior questionnaire data was checked against transcripts for high school grades, amount of coursework taken from the beginning of 10th grade through the end of the current school year in each of seven subjects, and whether or not each of six specific courses ever had been taken.

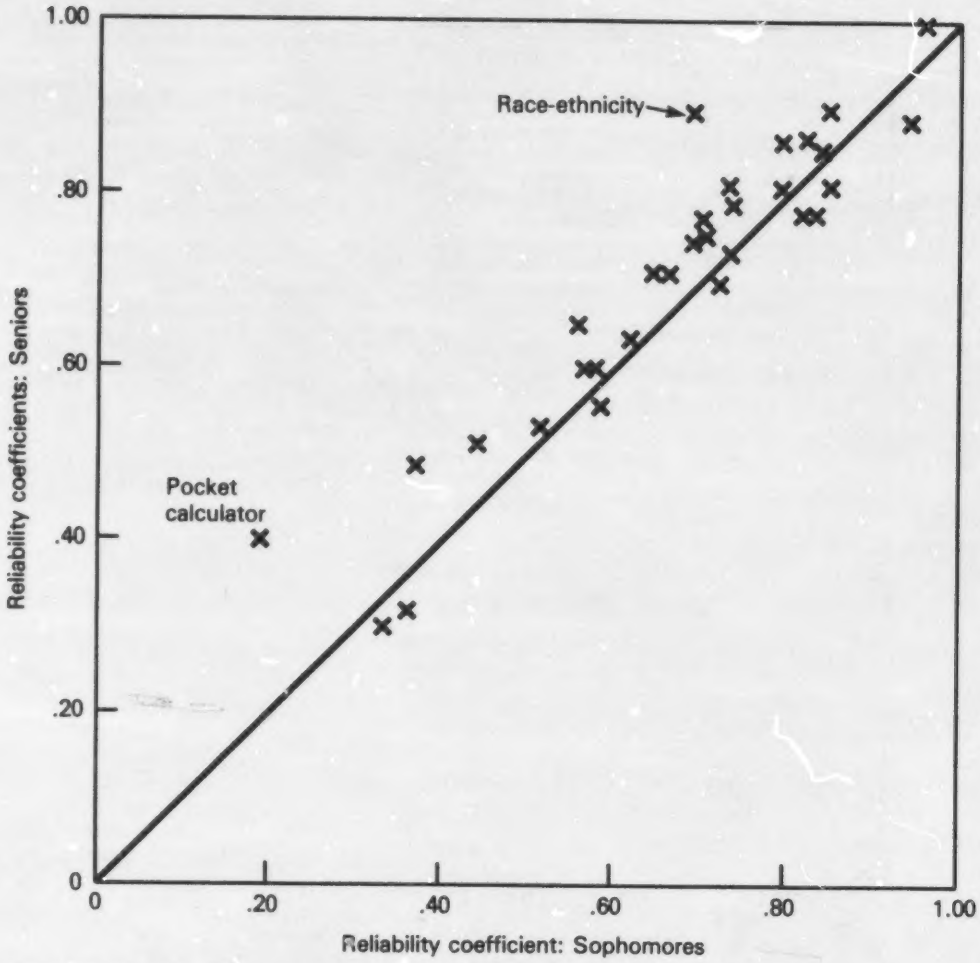
Grade point averages were calculated by NCES from the data on course grades and credits earned (Carnegie units) in the transcript file. Grades had been calibrated to a common scale by NORC based on a standardized 13-point grade scale.[26] Students reported grade averages on an 8-point scale ranging from "Mostly A (a numerical average of 90-100)" to "Mostly below D (below 60)." The numerical values associated with letter grades in the student questionnaire did not agree with the corresponding numerical values for any of the three

[24]Conger *et al.*, (1976) also concluded based on a review of reliability and validity studies that validity results generally are in agreement with the findings on reliability, with the reliability level putting a constraint on the maximum value of the validity coefficient.

[25]The only two items shown in figure 3 to have reliability coefficients more than marginally lower than their validity coefficients (namely, sophomore pocket calculator and race/ethnicity items) are ones for which the sophomore reliability coefficient was unusually low relative to the senior reliability coefficient (figure 2).

[26]Code values ranged from 1 for A+, 2 for A-, 3 for B+, down to 13 for F. Numerical grade values (e.g., 74, 93) were assigned code values using one of three grade scales depending on the minimum passing grade. See Jones *et al.* (1983).

Figure 2.—Estimated reliability coefficients for student responses to family background questionnaire items: 1980 senior vs. 1980 sophomores*

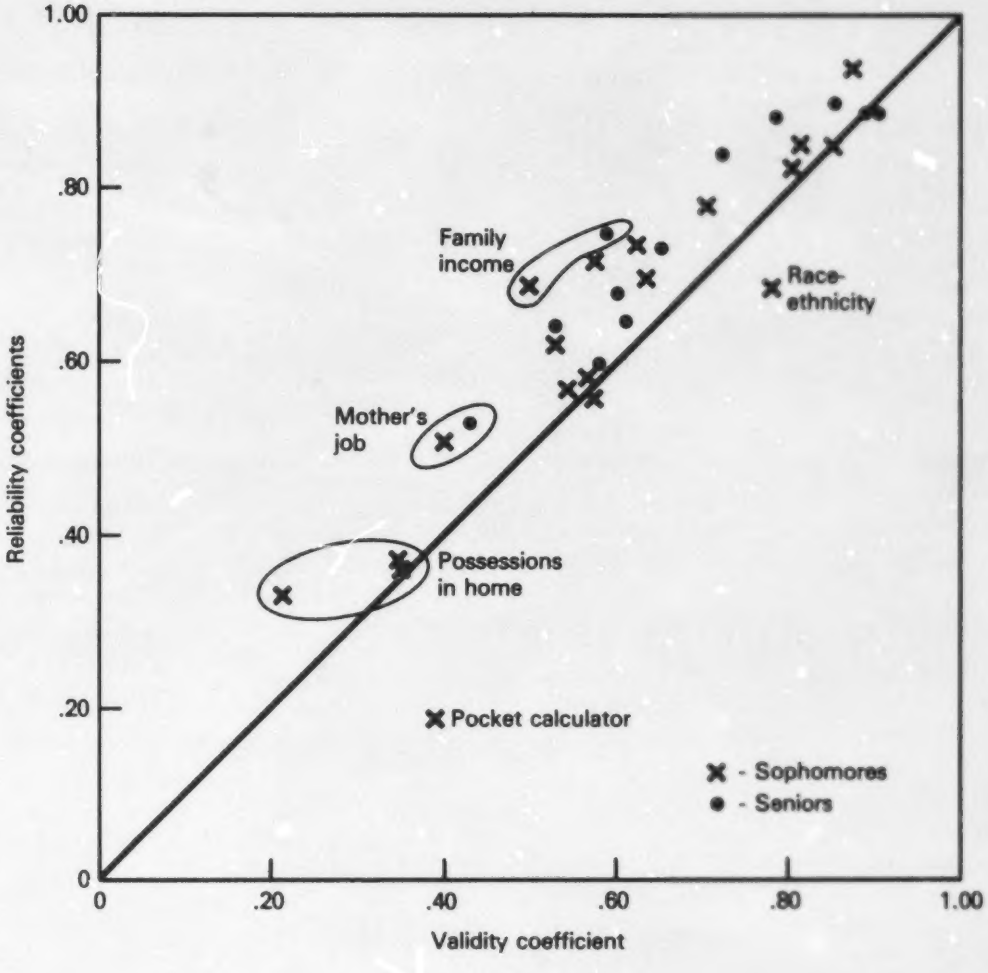


*See tables 3.3. and A-3.

grade scales used in coding transcripts.[27] Despite such problems, the responses of students correlated fairly highly (.77) with grade averages derived from their transcripts (table A.4). This coefficient is not quite as high as one might have expected based on the correlation coefficient of .76 between senior and parent data found in a subsequent section of this report. The parent questionnaire response categories, however, were identical with and thus directly comparable to those in the student questionnaire, whereas calculation of grade point averages from high school transcripts involved a number of assumptions.

[27]For example, the student questionnaire category was "Mostly B (80-84)," whereas in the transcript coding scheme, B corresponded to numerical values of 83-86, 86-88, or 88-91, depending on the minimum passing grade.

Figure 3.—Comparison of estimated reliability and validity coefficients for student responses to family background questionnaire items: 1980 sophomores and seniors



Seniors were asked, "Starting with the beginning of the *tenth* grade and through the end of this school year, how much coursework will you have taken in each of the following subjects? Count only courses that meet at least three times (or three periods) a week." Eight answer categories were provided ranging from "None" to "More than 3 years" in half-year intervals. For this study, transcript data on total number of credits earned for the three school years 1979-80, 1980-81, and 1981-82 were calculated and mapped into the eight student answer categories.[28] This was done for the subject areas that correspond to the "new basics," as described by the National Commission on Excellence in Education in its 1983 report *A Nation At Risk: The Imperative for Educational Reform*. Remedial English and mathematics courses were included in the transcript counts of units earned.[29]

[28]Transcript course titles were coded in accordance with the taxonomy given in Ludwig *et al.* (1982).

[29]Further information on the assignment of courses to the subject areas examined in this report is available on request from NCES.

The quality of student reports on amount of coursework was found to vary considerably by subject area: from highs of about .87 for French, German, and Spanish; to .70 and .66 for science and mathematics, respectively; to .39 for history or social studies and .28 for English or literature (table A.4). It should be noted that all students, regardless of their academic skills and other characteristics, tend to take about the same amount of coursework in the areas that have the lowest coefficients; but students vary considerably more in the amount of coursework taken in areas that have high coefficients. For example, students in the top test score quartile took only one-quarter more semester in English and literature but over more semesters of science than those in the bottom quartile (table A.5). Furthermore, the standard deviations of the distributions for all seniors were only 1.1 semesters for English but 1.9 semesters for science (not tabled).

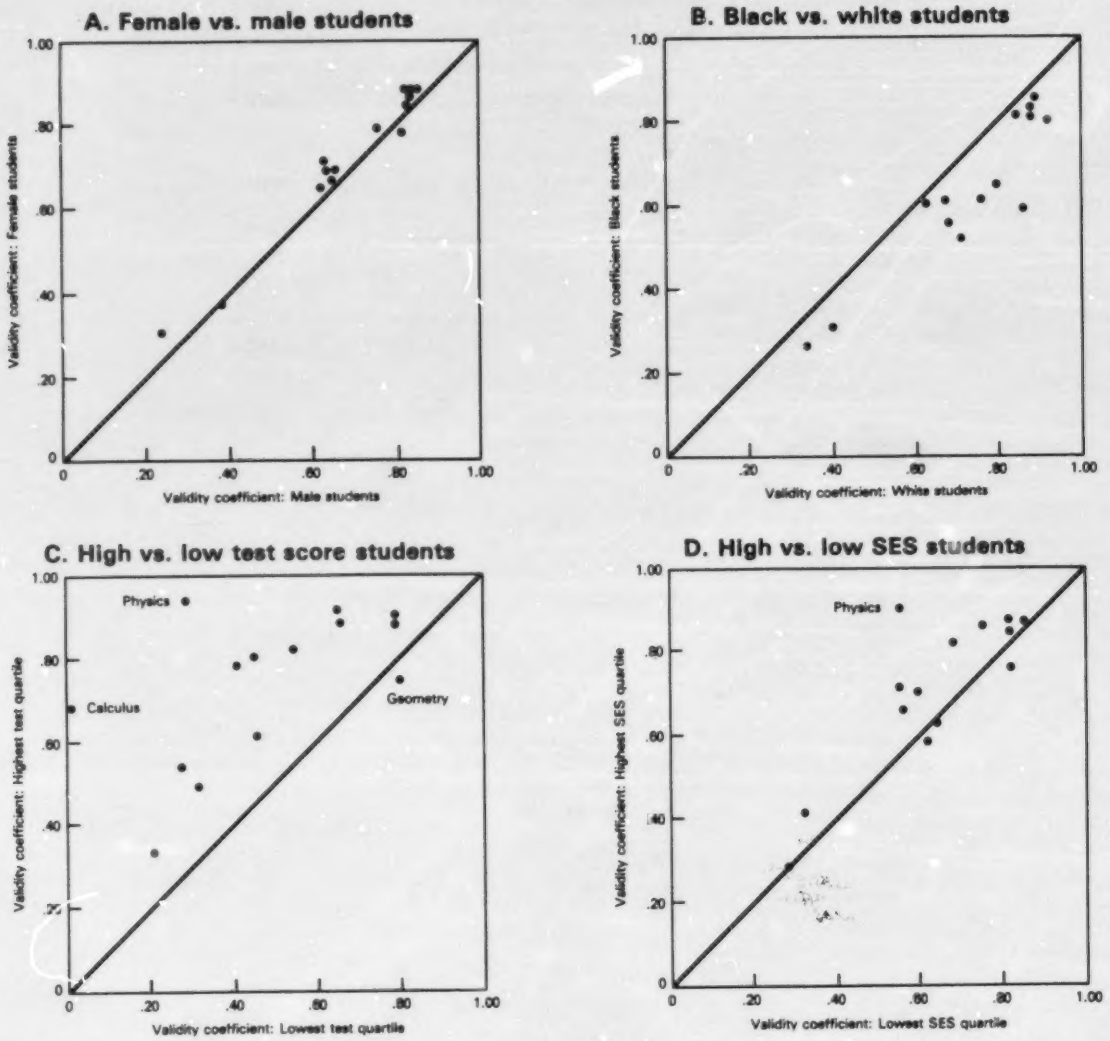
The third, but related, area of data quality investigated by use of transcript data was whether the senior ever had taken (counting the present semester) each of six specific courses in mathematics and science. The student was considered to have taken a course in, say, trigonometry if a unified mathematics course was taken that included trigonometry. Coefficients were found to be in the .80's for chemistry (.87), physics (.80), and geometry (.85) and in the .60's for second-year algebra (.68), calculus (.67), and trigonometry (.63) (table A.4). To the extent that some students might have taken some of these courses earlier than the 9th grade, these coefficients would be underestimates of the quality of student reports. As will be seen shortly, there is some evidence that this is the case for second-year algebra, geometry, and trigonometry. Thus, again because of use of a somewhat fallible or noncomparable standard, the true quality of student reports of coursework in these areas undoubtedly is somewhat higher than indicated by the coefficients estimated in this study.

The quality of student information on grades and coursework is not the same for all students. Subgroups differ very much in the same way as they were found to differ with regard to validity of family background variables. Data quality is slightly higher for females than for males; and it is considerably higher for white students, students who score high on cognitive tests, and students from high SES backgrounds than it is for Hispanics or blacks, students who do poorly on cognitive tests, and students from low SES backgrounds (tables 3.5 and A.4 and figure 4). The relation of data quality to test score quartile is especially strong, the mean coefficient varying from .47 to .73. The fact that an otherwise very strong pattern is broken between the third and top quartiles for three of the mathematics courses (table A-4) probably reflects a problem with the standard. Perhaps the brightest students took some of their mathematics courses prior to the period covered by the transcripts.

Table 3.5. Mean validity coefficients for 14 school-related variables for subgroups of 1980 sophomores in 1982 (transcript study)

Subgroup	Mean validity coefficient
Sex:	
Male	.689
Female	.721
Race/ethnicity:	
White	.734
Hispanic	.600
Black	.630
Test score quartile:	
Lowest	.472
Second	.620
Third	.706
Highest	.734
Socioeconomic status quartile:	
Lowest	.633
Second and third	.706
Highest	.707

Figure 4.—Estimated validity coefficients for student responses to items about grades and coursework (transcript data as standard): 1980 sophomores in school in 1982*



*See tables 3.3 and A.4.

3.2.2 Bias—transcript as standard. To examine the bias aspect of the quality of student responses for grades, it was first necessary to map the 13-point transcript scale into the 8-point student scale.[30] (Because the highest interval was open-ended, the amount of coursework already had been mapped into the student questionnaire categories for the correlational aspect of the investigation of data quality.) A further transformation then was made to the data from both sources so that A = 4 (previously coded 1), B = 3 (previously coded 3), C = 2 (previously coded 5), and D = 1 (previously coded 7).

[30]Transcript grade-point average values were transformed to student questionnaire code values as follows:

Transcript	Student
Less than 2.50	1 (Mostly A's)
2.50 to 4.49	2 (About half A's and half B's)
4.50 to 5.49	3 (Mostly B's)
5.50 to 7.49	4 (About half B's and half C's)
7.50 to 8.49	5 (Mostly C's)
8.50 to 10.49	6 (About half C's and half D's)
10.50 to 11.49	7 (Mostly D's)
11.50 or higher	8 (Mostly below D)

Students tended to report slightly higher grade averages than those calculated from their transcripts (table 3.6). The student mean is about one-fifth of a letter grade above the transcript mean. The amount of bias is slightly higher for black than for white or Hispanic students (0.3 vs. 0.2 letter), but the variation in bias values among sex, test score, and SES groups is negligible (table A.5).

Students also tended to overstate the amount of coursework they had taken or were taking in all areas except history or social studies (table 3.6). The size of the bias is greatest for mathematics (1.1 semester), next greatest for science (0.6 semester), and one-fourth a semester or less for each of the other subject areas.

Table 3.6. Estimated bias in responses of 1980 sophomores to 1982 questionnaire items about grades, coursework, and courses

Subgroup	Source		Bias
	Student	Transcript	
Grade average (letter grade units)*	2.84	2.62	.22
Amount of coursework (years):			
Mathematics	4.15	3.07	1.08
English or literature	5.81	5.56	.25
French	.58	.44	.14
German	.17	.13	.04
Spanish	.94	.72	.22
History or social studies	4.58	4.63	-.05
Science	3.43	2.87	.56
Ever have taken or are taking (percentage):			
Second-year algebra	48.7	38.1	10.6
Geometry	55.1	48.8	6.3
Trigonometry	25.6	14.4	11.2
Calculus	8.8	5.3	3.5
Physics	19.5	15.3	4.2
Chemistry	36.9	32.2	4.7

*Scaled A = 4, B = 3, C = 2, D = 1.

The magnitude of the bias in amount of coursework is about the same for the various subgroups of students for two of the four areas for which subgroup comparisons were made (viz., English or literature and history or social studies); but for mathematics and science, bias is considerably higher for Hispanics, blacks, students who score low on cognitive tests, and students from low SES backgrounds than it is for other students (table A.5).[31] For example, the bias in student reports of amount of mathematics is 1.8 semesters for black students, but only 0.9 semester for white students. In fact, the mean self-reported figure is higher for black than for white students (4.5 vs. 4.2 semesters), whereas the transcript mean is lower (2.7 vs. 3.3). To cite a second example, the bias for mathematics, proceeding from lowest to highest test score quartile, declines from 1.8 to 0.6 semester.

Finally, it was found that students tended to over-report that they had taken (or were taking) specific mathematics and science courses (table 3.6). For example, student questionnaires yield an estimate of 55 percent for geometry as compared with an estimate of 49 percent obtained from transcripts.[32]

3.2.3 Reliability—twin file. Table 3.7 summarizes the information contained in table A.6 on reliability of school-related variables. This is based on the extent of agreement of the questionnaire responses from the members of twin sets. The coefficients cover a very wide range of values—from less than .10 to over .90.

The coefficients are fairly high for school attendance and student-body composition variables. The reliability of student responses to the question about whether the school has a minimum competency or proficiency test

[31] This part of the analysis was limited to the four major subject areas that are mentioned.

[32] Subgroups were not compared with regard to bias in student reports of whether specific courses had been (or were being) taken.

program is disappointingly low. The school questionnaire, however, is a more authoritative source of information for this variable; and there would seem to be little if any need to resort to student-reported data.[33]

Table 3.7. Summary of reliability coefficients for school-related variables (twin file): 1980 sophomores and seniors

Variable	Reliability coefficient	
	Sophomores	Seniors
Kindergarten attendance	.93	–
When began attending this school	–	.89
Number of school changes	.77	–
School racial/ethnic composition, earlier grades	.58 to .76	.68 to .83
Minimum competency test	.45	.68
Enforced school rules	.26 to .54	–
School disciplinary problems	.18 to .38	–
School rating	.21 to .39	.24 to .52
Instructional methods	–	.05 to .33
School problems interfering with education	–	.14 to .24

As would be expected, the coefficients associated with the remaining five items are moderate or low in size. It should be noted that the two members of each twin pair were not necessarily exposed to identical school environments and experiences. Thus, to the extent they were not, the reliability coefficients for some items (such as instructional methods experienced) may be underestimated. On the other hand, the responses of twins will agree more closely than will the responses of two randomly selected students within a school to the extent that twin responses reflect that influence of attitudes and standards that are learned in the common home.[34] More importantly, two other considerations should be kept in mind for these last five items. First, each of them has five to eight sub-items. In many instances it is more appropriate to combine these sub-items into composite measures than to employ them as individual sub-items. Second, for many analyses it is appropriate to aggregate the data for these kinds of items to the school level. The reliability of a composite measure, of course, would be greater than that of its component sub-items; and the reliability of a mean of up to 36 observations would be considerably higher than the reliability value for an individual student.[35]

In figure 5, the reliability coefficients for seniors are plotted against those for sophomores. Seniors provided considerably more reliable school data than did sophomores. For the 15 variables plotted in figure 5, the mean reliability coefficients are .537 for seniors, .439 for sophomores.

3.3 Other Variables

This section of the report examines the consistency of parent-child responses to questions primarily regarding perceived influences, expectations, aspirations, and attitudes. In only one instance (namely mother's aspirations for child's education) can the statistics be considered to measure the quality of student questionnaire data.

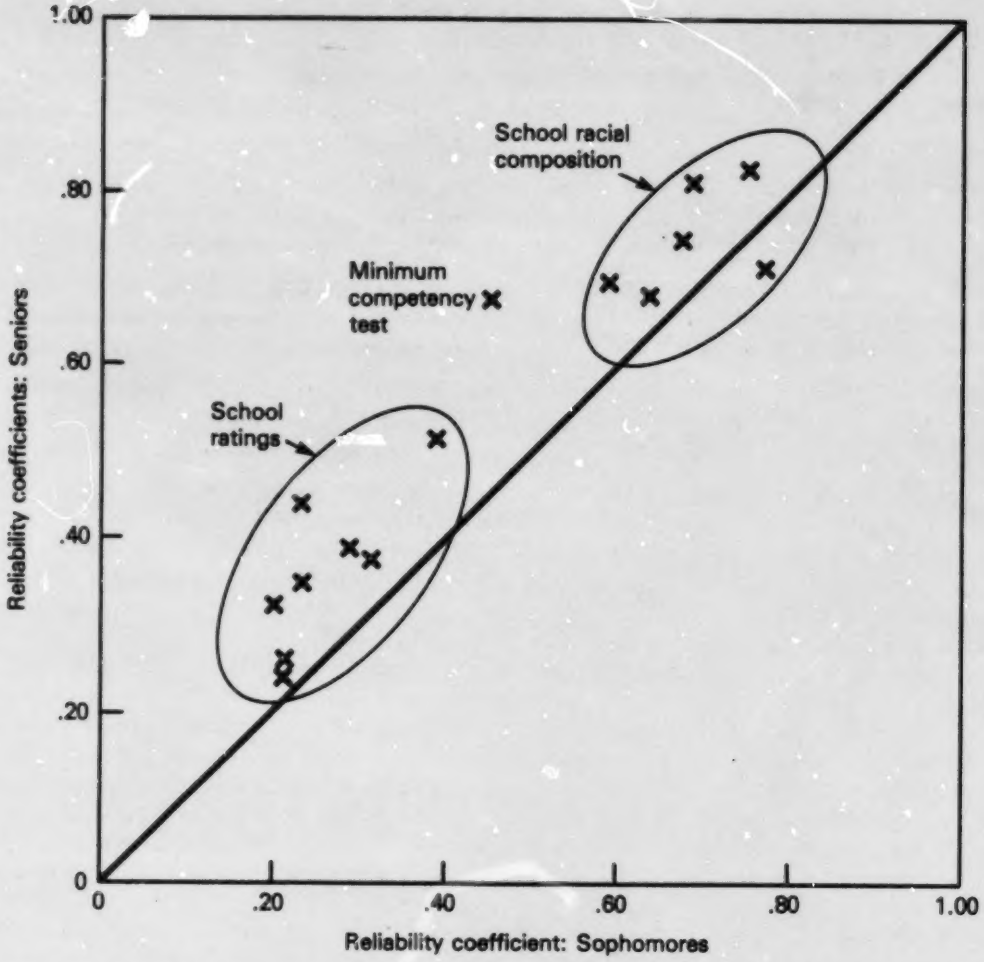
3.3.1 Consistency of parent-child responses—grades, plans, attitudes. Table A.7 shows the estimated amount of correlation between parent and child responses to questions concerning the child's grades, plans, and attitudes. The results are summarized in Table 3.8.

[33]The validity of student reports was not checked against minimum competency data reported in the school questionnaire. The validity coefficients, however, should be slightly lower than the reliability coefficients (section 3.1.4).

[34]Some preliminary analysis by NCES indicates that the correlation between twin responses tends to be about 50 percent higher than that for randomly selected students in the same school for the questionnaire items pertaining to school discipline problems, school rules, and school ratings.

[35]For example, if the reliability of a variable is .20, the reliability of the mean value of the variable for 36 randomly selected students is estimated by the Spearman-Brown formula to be .90. (Lord and Novick, 1968). It should be noted, however, that the responses of twins will tend to agree more closely than will the responses of two randomly selected individuals within a school. Thus the Spearman-Brown formula will tend to over-estimate the reliability of school means derived from twin-based estimates of correlation coefficients.

Figure 5.—Estimated reliability coefficients for student responses to school-related questionnaire items: 1980 sophomores and seniors*



*See tables 3.3 and A-8.

Table 3.8. Summary of consistency coefficients—Grades, plans and attitudes

Variable	Consistency coefficient	
	Sophomores	Seniors
High coefficient (over .70):		
High school grade average	.75	.76
Moderate coefficient (.50 to .69):		
Mother's aspirations for child's education	.57	.59
Low coefficients (.00 to .49):		
Age expects (is expected) to marry, etc.	.27 to .43	.40 to .57
College expectations in earlier grades	.43 to .45	.40 to .46
Child's ability to complete college	.40	.42
Child's occupational expectations	.31	.37
Influence of parents on post-high school plans	-	.18, .21
Sex role attitudes	.18 to .21	-

The amount of agreement between parent and child is highest for grade average (about .76)[36] and for mother's aspiration for child's education (about .58). There is moderately good correspondence in answers between seniors and their parents regarding the age at which certain events are expected to occur: marriage (.57), first child (.51), start regular job (.48). All other coefficients are less than .48, and some are as low as about .20 (namely, those for parental influence on post-high school plans and sex role attitudes).

Figure 6-A reveals that the answers of seniors and their parents agree more closely than do the answers of sophomores and their parents. The mean coefficients for 10 variables are .498 and .431, respectively.

There was better agreement between daughters and parents than between sons and parents regarding events related to family formation: marriage, first child, and home ownership (table A.7 and figure 6-B). Otherwise, the estimated coefficients for sons and daughters are about the same.

For most variables consistency coefficients are lower, often considerably so, for minority-group students than for white students (table A.7 and figure 6-C). The average coefficient for 28 variables is .300 for black students, .375 for Hispanic students, and .436 for white students.

Consistency between parent-child responses regarding high school grade average is higher for students who scored well on tests than for those who made low scores—about .78 vs. about .58 (table A.7). Otherwise, the points plotted in Figure 6-D indicate no clear edge to students who made high test scores.[37]

The pattern of results by SES quartile is mixed (figure 6-E). There is better parent-child agreement for high than for low SES students for grades, sex role attitudes, and age the student will get a regular job, but not for most other variables. The results are particularly favorable to low SES students for the variables concerning when the student will marry, have first child, and live in own home, and whether the student had college plans when in earlier grades.

It makes little difference which parent answers the parent questionnaire (figure 6-F). The mean consistency coefficients are .40 for both mother-child and father-child contrasts.

3.3.2 Mean difference between parent and child responses—grades, plans, attitudes. Using the child's report as a basis of comparison, table A.8 shows that: (a) parents were more likely to report high grade averages (40 vs. 35 percent B+ or better, seniors); (b) parents were more likely to believe events in the lives of their children will occur later (e.g., 33 percent of their parents but only 27 percent of sohomores thought marriage will occur at age 25 or later); (c) mothers were more likely to report that the mother expects the child to get some, but less than 2 years of postsecondary education (5 vs. 19 percent for sophomores); (d) parents were more likely to believe "yes, definitely" the child has the ability to complete college (58 vs. 39 percent, sophomores), but were less likely to report that they expected the child to go to college when the child was in earlier grades (69 vs. 75 percent for 8th and 9th grades, sophomores).[38]

3.3.3 Consistency of parent-child responses—financial and college items. The highest correlation coefficient between parent and child responses to questions regarding living and schooling expenses, college costs, and college choice factors is .61 for estimated schooling expenses, September 1980 through August 1981 (table A.9). The coefficient for estimated living expenses for the same period, on the other hand, is quite low (.22).

The consistency coefficients are low (.07 to .38) for 11 of 13 possible sources of money to pay for the student's living and schooling expenses. The two exceptions are "parents" and "social security or VA (Veterans Administration) benefits." The estimated coefficient for each of these two sources is .47.

Parent and child estimates of the cost of attending various kinds of postsecondary institutions are only weakly correlated. The coefficients range from .06 to .25.

Finally, parents and children assessed the importance of seven factors in choosing a college. The highest consistency coefficients are for "able to live at home and attend college" (.56) and "availability of financial aid" (.44). The other five coefficients vary from .11 to .32.

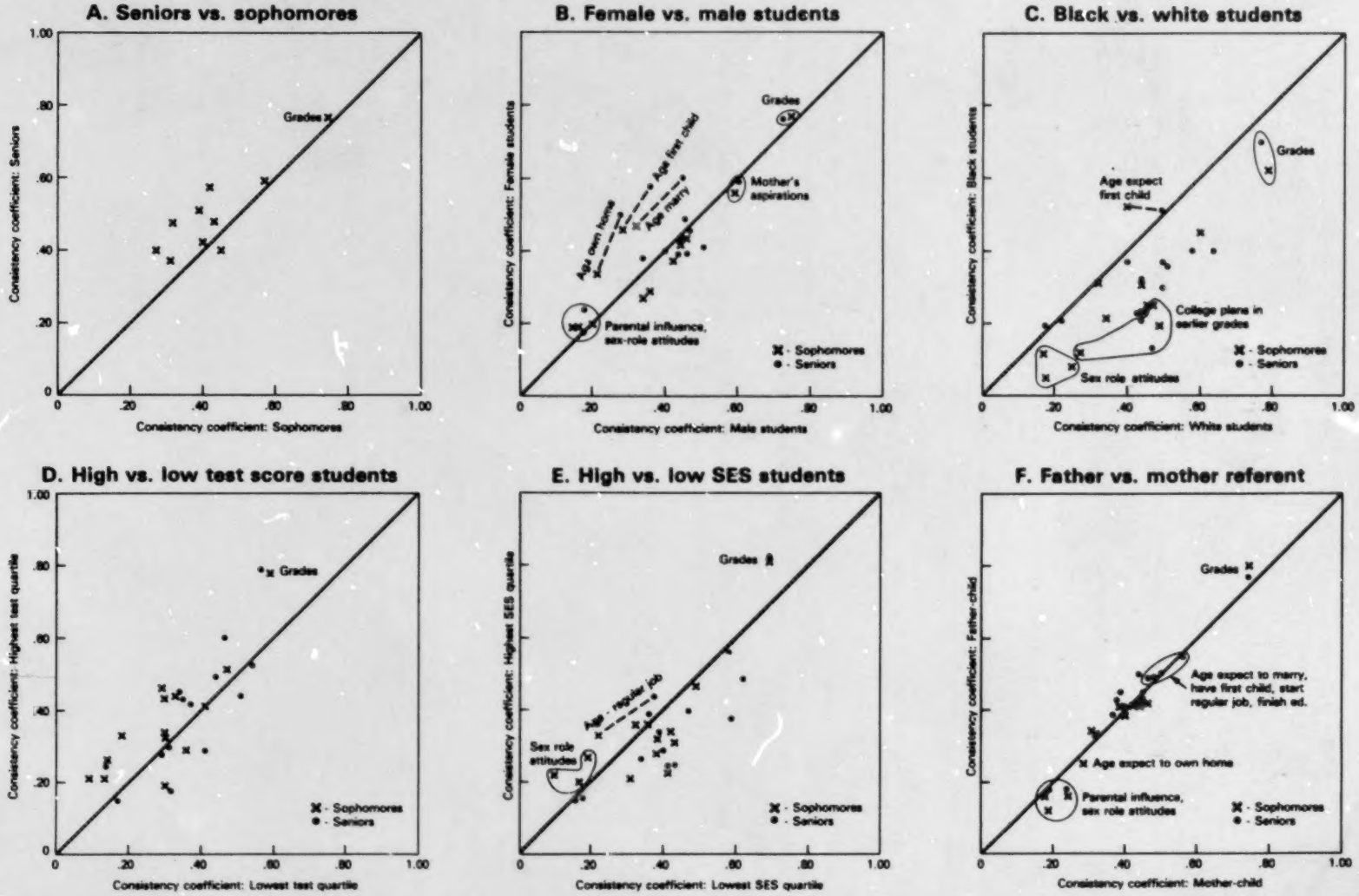
---

[36]Since the parent's report does not set a factual standard, .76 may be considered a lower bound for the validity of student self-reports of grade point averages.

[37]The number of points above the line in Figure 6-D is not significantly different from the number below. Excluding grade average, the mean coefficient is .355 for high-scoring and .315 for low-scoring students.

[38]All of the estimates of relative bias cited in this section and a later section concerning financial college items differ from zero by at least two standard errors. (See table A.12 for standard error values.)

Figure 6.—Estimated coefficients of consistency between child and parent responses to questionnaire items about grades, attitudes, and plans*



*See tables 3.3 and A-7.

BEST COPY AVAILABLE

As may be observed in figures 7-A through 7-E, there is greater consistency in parent-child responses for some kinds of students than for others. The mean coefficient for the 28 variables is higher for female than for male students (.267 vs. .229), for white (.267) than for Hispanic (.180) or black (.171) students, for students who scored high on cognitive tests than for those who scored low (.300 vs. .159), and for students from high SES backgrounds than for those from low SES backgrounds (.279 vs. .200). Finally, the responses of students were more consistent with those of their fathers than those of their mothers (.272 vs. .232).

3.3.4 Mean difference between parent and child responses—financial and college items. Parents were much more likely than seniors to report that the child would not require any money for living expenses for the period September 1980 through August 1981, (42 vs. 27 percent) (table A-10). They also were more likely to report that the child would not require any money for schooling expenses during the same period (28 vs. 23 percent). As illustrated in table 3.9, parents estimated the yearly costs of attending postsecondary education institutions to be lower than the costs estimated by their children.

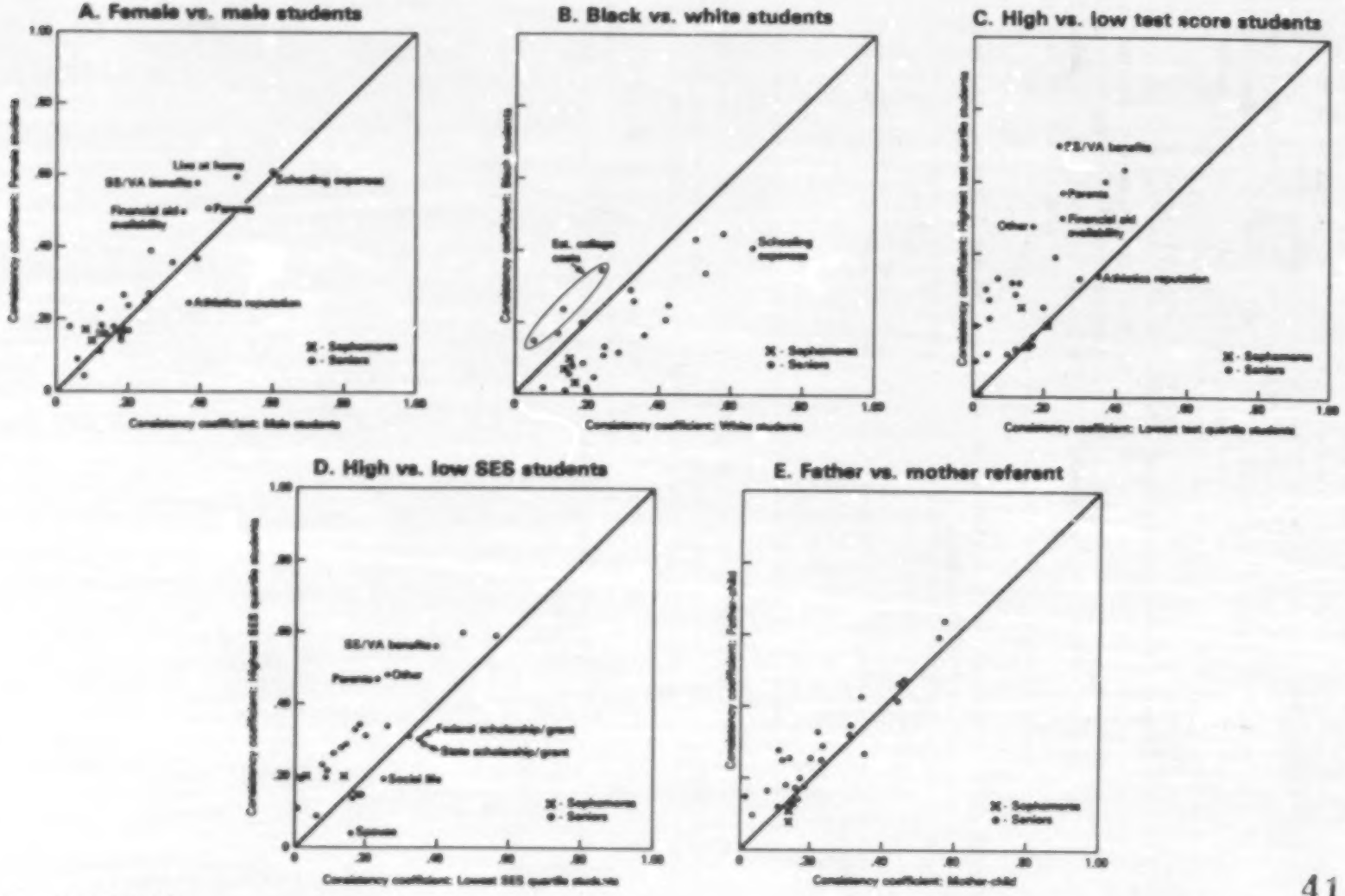
Table 3.9. Relative bias in estimates by sophomores and their parents of cost of attending postsecondary institutions, by type of institution

Type of Institution	Est. yearly cost	Percent	
		Sophomores	Parents
Junior or community college	Less than $1,000	50	59
State 4-yr college or university	Less than $2,000	30	42
Private 4-yr college or university	Less than $5,000	50	56

Seniors were considerably more optimistic than their parents that they would be able to pay for at least some of their living and school expenses from their own savings (71 vs. 35 percent) and earnings (74 vs. 63 percent) and by means of a scholarship, grant, or loan (e.g., 26 vs. 17 percent for Federal scholarship or grant).

Finally, parents were more likely than seniors to rate as "very important" in choosing a college: "college expenses" (53 vs. 37 percent;)"reputation in academic areas" (75 vs. 53 percent); and "able to live at home and attend college" (33 vs. 22 percent); but parents were less likely to consider the "social life at the college" to be very important (11 vs. 28 percent).

Figure 7.—Estimated coefficients of consistency between child and parent responses to financial and college-related questionnaire items: 1980 sophomores and seniors*



*See tables 3.3 and A-9.

BEST COPY AVAILABLE

Chapter 4 — Summary and Discussion

This study employed data collected from parents, high school transcripts, and twins to examine the validity and reliability of responses of high school students (sophomores and seniors) to questionnaires administered to them in a group setting. The validity portions of the study examined a variety of family background and home environment variables (23 items), high school grade average, amount of coursework taken (7 items), and whether certain specific courses had been taken (6 items). Data collected from twins, the reliability portion of the study, broadened the scope of the investigation of data quality. Additional family background and home environment variables could be examined with twin data (32 items in all plus an SES composite), as well as variables concerning school conditions, problems, and practices (20 items) and the students' earlier and current school experiences (17 items). The consistency of child and parent responses to questions dealing with aspirations, expectations, attitudes, opinions and plans also was examined (43 items).

The results of these examinations of data quality were presented in the previous chapter. They have a number of implications for the analysis and interpretation of HS&B data and for the design of future surveys.

A. Data collection.

While highly reliable data can be obtained for some events that happened in the past (e.g., kindergarten attendance), there generally is a decay in quality as the recall period lengthens (e.g., with regard to whether the student's mother worked at various stages of the student's life). This result is consistent with the findings of other researchers (section 1.2); and suggests, for example, that adults would not provide very accurate information about many aspects of their earlier high school days.

The difficulty of respondents to accurately recall past events and experiences, much less earlier attitudes, of course, is the primary reason for undertaking prospective longitudinal studies such as HS&B. Longitudinal studies are essential to collect reliable event history and other data. Cross-sectional surveys cannot be relied upon to gather accurate data about the past.

B. Composites and means.

The vast majority of the variables examined in the study were contemporaneous rather than retrospective in nature; and most of these dealt with factual rather than attitudinal matters. The validity and reliability coefficients for these contemporaneous, factual variables exhibited considerable variation, from about .20 to over .90. The only home and family background variables that had coefficients less than .50 for both cohorts were four that can enter a "household possessions" composite, which in turn can be employed as components of an SES composite. It is very encouraging that the SES composite that often is employed in analyzing HS&B data not only has a high coefficient (mid .80s), but is available for virtually all students.[39]

Quite a few of the school-related variables required the student to exercise judgment, for example, in forming an opinion regarding the extent to which certain school rules were enforced. The validity (or reliability) values for these opinion variables tended to be lower than for the more factually oriented ones such as whether certain specific courses had been taken. The opinion variables, however, generally are quite suitable for use in composites.

It seems quite clear from these results that analysts should make full and appropriate use of composites (such as possessions in home, SES, and several that may be formed from items regarding school climate). Not only should the composites have high validity and reliability coefficients, but values of the composites should be available for virtually all subjects. On the other hand, use of selected individual component items as proxies for the entire set generally should be avoided.

[39] A composite value was calculated for students who provided data for at least two of the five components. Thus about 95 percent of the students have SES composite values even though data are available for only 80 percent of the students for some component items.

In several instances, students were used as observers and reporters of school conditions, policies, and practices. Not only should the data analysts consider the use of composites in these instances but they also should consider whether it would be more appropriate to aggregate the data to the school level and employ school averages rather than individual student values in the analysis. The reliability of the average, of course, would be considerably higher than the reliability value for an individual student.

C. Statistical adjustment and data interpretation.

Errors of measurement in either or both variables attenuate or weaken correlation coefficients and, if the error is in the independent variable, attenuate regression coefficients. The validity (or reliability) coefficients presented in this report may be employed to adjust for this attenuation or may be employed in models that explicitly incorporate a provision for measurement error. Validity coefficients,[40] however, were found to depend on the characteristics of the students answering the question. The coefficients tended to be slightly higher for seniors and females than for sophomores and males but they were considerably higher for students who scored high on cognitive tests, high SES students (school variables),[41] and white students than for those who performed poorly on the tests, low SES students (school variables), and Hispanic and black students. Thus it is important that the adjustment factors used by the analyst be the ones that correspond to the subgroup under investigation.

For some school-related variables, data provided by these subgroups differed not only in the degree of correlation with the standard, but also in the size of the bias of the response. For example, the correlation coefficient for students in the top and bottom test score quartiles for number of semesters of mathematics taken by the student were .78 and .41, respectively; and the same two groups overestimated the amount of mathematics taken by an average of 0.6 and 1.8 semesters, respectively. Thus the use of student reports could result in a serious underestimate of the difference between these two groups in the amount of mathematics taken and of the relationship of amount of mathematics taken to achievement.

The variation of validity coefficients and bias values among subgroups also means that the measurement error may well be correlated with the value of the variable itself and perhaps with values of some of the other variables in the model. Thus the assumption of "well-behaved" error employed in certain structural equation models will not be met.

It may be quite important for correct data analysis and interpretation to employ the transcript file for information about grades and coursework. This is particularly true for coursework in mathematics and science because of the large subgroup differences in bias values as well as validity coefficients, but it also is true for English or literature and history or social studies because their overall validity coefficients are so low (less than .40).

D. Future surveys.

Designers of student questionnaires for new studies should examine the wording and formatting of HS&B items, in conjunction with the values of the reliability and validity coefficients, and the item response rates. Attempts should be made to reduce the ambiguity and improve the clarity of items with low coefficients.[42] The special difficulty that certain subgroups may have in correctly understanding the item should be kept in mind during this review process. If there seems to be little potential for improvement of an item, deletion of the item should be considered. This consideration, of course, should take into account the importance and use of the item. For example, it is doubtful that one would want to eliminate family income because of its importance but a "home possession" item with the same (or even higher) validity coefficient as the income item might be deleted with little harm.

Although the quality of data provided by most students about grades and coursework generally is quite good, it is substantially poorer for low achieving and certain other subgroups of students. These same subgroups also tend to overstate to a greater degree than other students the amount of coursework taken in some subject areas. This situation could easily lead to the formation of erroneous conclusions. Thus it is important that sufficient resources be allocated in future surveys to the collection and processing of the high school transcripts of the sampled students.

---

[40] The sample size for twins was not large enough to support a similar investigation of reliability coefficients except for a comparison of those of sophomores and seniors.

[41] The results for home and family background variables are uncertain because many of these variables entered into the composite utilized to classify students by SES.

[42] The items which had coefficients less than .60 are: six home possession items (place to study, encyclopedia or other reference books, more than 50 books, pocket calculator, daily newspaper, and two or more cars or trucks that run); mother's occupation; house owned or rented; mother worked before child attended elementary school; mother lives in same household; amount of coursework in English or literature and history or social studies; and the items contained in the five questions dealing with school problems, practices, and conditions.

References

Astin, A. W. *Who Goes Where to College?* Chicago: Science Research Associates, 1965.

Averich, Harvey A.; Carroll, Stephen J.; Donaldson, Theodore S.; Kiesling, Herbert J.; and Pincus, John. *How Effective Is Schooling?* Rand Corp., Santa Monica, 1971.

Bishop, Yvonne M. M.; Fienberg, Stephen E.; and Holland, Paul W. *Discrete Multivariate Analysis: Theory and Practice.* Cambridge, Mass.: MTI Press, 1975.

Boruch, Robert F. and Creager, John A. "Measurement error in social and educational survey research." *ACE Research Reports,* 7(2), 1972.

Borus, Michael E. and Nestel, Gilbert. "Response bias in reports of father's education." *Proceedings of the Social Statistics Section,* 1971, Washington, D.C.: American Statistical Association, 1972, 263-268.

Borus, Michael E. and Nestel, Gilbert. "Response bias in reports of father's education and socioeconomic status." *Journal of the American Statistical Association, 1973, 68,* 816-820.

Cannell, Charles F. and Fowler, Floyd J. "Comparison of a self-enumerative procedure and a personal interview: A validity study." *Public Opinion Quarterly,* Summer 1963, *27,* 250-264.

Cohen, Roberta S. and Orum, Anthony M. "Parent-child consensus on socioeconomic data obtained from sample surveys." *Public Opinion Quarterly,* Spring 1972, *36,* 95-98.

Conger, Anthony J.; Conger, Judith C.; and Riccobono, John A. *Reliability and Validity of National Longitudinal Study Measures: An empirical reliability analysis of selected data and a review of the literature on the validity and reliability of survey research questionnaires.* Report prepared by the Research Triangle Institute for the National Center for Education Statistics, 1976.

Dresch, Stephen P.; Stowe, Peter S.; and Waldenberg, Adair L. *Parental Capacity to Finance Their Child's Postsecondary Education.* Report prepared by the Institute for Demographic and Economic Studies for the National Center for Education Statistics, in preparation.

Echternacht, Gary J. "The response stability of the base-year National Longitudinal Study data" and "The validity of the base-year National Longitudinal Study data," Appendix F of *The Base-Year Survey of the National Longitudinal Study of the High School Class of 1972.* Report prepared by the Educational Testing Service for the National Center for Education Statistics, 1973.

Frankel, Martin R.; Kohnke, Luanne; Buonanno, David; and Tourangeau, Roger. *Sample Design Report.* Report prepared by the National Opinion Research Center for the National Center for Education Statistics, 1981.

Hansen, Morris H.; Hurwitz, William N.; and Madow, William G. *Sample Survey Methods and Theory,* Vol. 1. New York: John Wiley, 1953.

Heyns, Barbara and Hilton, Thomas. "The cognitive tests for High School and Beyond: An assessment," *Sociology of Education, 55,* 89-102.

Jones, Calvin; Knight, Shirley; Butz, Marjorie; Crawford, Ioanna; and Stephenson, Bruce. *High School and Beyond: Transcript Survey (1982) Data File User's Manual.* Report prepared by National Opinion Research Center for the National Center for Education Statistics, 1983.

Kendall, Maurice G. and Stuart, Allen. *The Advanced Theory of Statistics,* Vol. 1. New York: Hafner, 1958.

Kendall, Maurice G. and Stuart, Allen. *The Advanced Theory of Statistics,* Vol. 2. New York: MacMillan, 1979.

Kerckhoff, Alan C.; Mason, William M.;and Poss, Sharon Sandomirsky. "On the accuracy of children's reports of family social status." *Sociology of Education,* Spring 1973, *46* 219-247.

Kish, Leslie. *Survey Sampling.* New York: John Wiley, 1965

Kyaser, Brian D. and Summers, Gene F. "The adequacy of student reports of parental SES characteristics." *Sociological Methods and Research,* 1973, *1* (3), 303-315.

Lord, Frederick M. and Novick, Melvin R. *Statistical Theories of Mental Test Scores.* Reading, Mass.: Addison-Wesley, 1968.

Ludwig, Meredith J.; Olivetti, L. James; Sandberg, Nancy; and Waite, Barbara J. *A Classification of Secondary School Courses.* Report prepared by Evaluation Technologies Incorporated for the National Center for Education Statistics, 1982.

Lyons, N. I. and Moore, R. P. *National Longitudinal Study of the High School Class of 1972: Reliability of Retrospective Data.* Report prepared by the Research Triangle Institute for the National Center for Education Statistics, 1976.

National Center for Education Statistics. *High School and Beyond: Twins and Siblings File Users' Manual,* 1982.

National Opinion Research Center, *High School and Beyond: Information for Users of Base Year (1980) Data.* Report prepared for the National Center for Education Statistics, 1980.

National Opinion Research Center. *High School and Beyond: Parent Questionnaire Codebook.* Report prepared for the National Center for Education Statistics, 1981.

Peng, Samuel S.; Fetters, William B.; and Kolstad, Andrew J. *A Capsule Description of High School Students,* NCES, 1981.

Riccobono, John; Henderson, Louise B.; Burkheimer, Graham J.; Place, Carol; and Levinsohn, Jay R. *National Longitudinal Study: Base Year (1972) through Fourth Follow-Up Survey (1979) Data File Users Manual,* Vol II, App. K. Report prepared by the Research Triangle Institute for the National Center for Education Statistics, 1981.

Siegel, Sidney. *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill, 1956.

Swineford, Frances. "Test analysis: National Longitudinal Study of the High School Class of 1972," Appendix D of *The Base-Year Survey of the National Longitudinal Study of the High School Class of 1972.* Report prepared by the Educational Testing Service for the National Center for Education Statistics, 1973.

Tourangeau, Roger; McWilliams, Harold; Jones, Calvin; Frankel, Martin F.; and O'Brien, Frank. *High School and Beyond: First Follow-Up (1982) Sample Design Report.* Report prepared by the National Opinion Research Center for the National Center for Education Statistics, in preparation.

van Es, J.C. and Wilkening, Eugene A. "Response stability in survey research: A cross-cultural comparison." *Rural Sociology,* June 1970, *35* (2), 191-205.

Walsh, T. C. and Buckholdt. *Accuracy of retrospectively reporting work status and occupation five years ago.* U.S. Bureau of the Census working paper, Economic Statistics Branch, Population Division, Vol. 15, No. 3, 1970.

Walsh, W. B. "Validity of the self-report." *Journal of Counseling Psychology,* 1967, *14* 18-23.

Werts, C. E.; Linn, R. L.; and Joreskog, K. G. "Quantifying unmeasured variables." *Measurement in the Social Sciences* (H. M. Blalock, Jr., Ed.) Chicago, Ill.: Aldine Pub. Co., 1974.

Appendix

Estimated Coefficients of Validity, Bias, Reliability, and Consistency by Item Type for Selected Student Subgroups: 1980 High School Sophomores and Seniors

Table A.1. Estimated validity coefficients for responses of 1980 sophomores and seniors to family background questionnaire items, by selected student characteristics and sex of parent questionnaire respondent

Variable	Cohort	All students	Sex		Race/ethnicity			Test score/quartile				SES composite			Respondent		n
			Male	Female	White	Hispanic	Black	1	2	3	4	Low	Med	High	Mother	Father	
1. SES composite	Sophomore	.80	.78	.82	.82	.74	.74	.67	.71	.79	.84	—	—	—	.78	.83	3,208
2. Items in home																	
a. Specific place for study	Sophomore	.21	.17	.24	.20	.24	.20	.18	.19	.21	.22	.15	.16	.16	.19	.22	2,958
b. Encyclopedia or other reference books	Sophomore	.35	.28	.40	.31	.40	.30	.36	.32	.33	.24	.34	.27	.10	.38	.27	2,999
c. More than 50 books	Sophomore	.35	.37	.34	.30	.37	.38	.28	.31	.34	.32	.29	.24	.09	.32	.37	2,976
d. Pocket calculator	Sophomore	.39	.37	.39	.34	.40	.35	.38	.30	.35	.32	.36	.25	.15	.39	.36	2,991
e. Daily newspaper	Sophomore	.54	.54	.54	.58	.44	.39	.42	.52	.56	.66	.48	.47	.61	.53	.54	2,991
f. Two or more cars or trucks that run	Sophomore	.57	.55	.59	.57	.46	.50	.51	.50	.56	.70	.51	.51	.63	.56	.56	2,984
g. Typewriter	Sophomore	.63	.63	.64	.65	.62	.56	.58	.58	.63	.69	.59	.57	.60	.62	.64	2,978
h. Electric dishwasher	Sophomore	.85	.84	.87	.87	.73	.72	.75	.82	.88	.90	.70	.84	.80	.84	.86	2,943
3. Parent's education																	
a. Mother	Sophomore	.81	.80	.83	.84	.70	.69	.72	.70	.83	.87	.61	.66	.80	.81	.82	2,630
Mother	Senior	.85	.84	.86	.87	.74	.78	.77	.77	.84	.91	.66	.74	.79	.83	.88	2,708
b. Father	Sophomore	.87	.86	.88	.88	.86	.70	.78	.80	.86	.90	.41	.70	.80	.86	.88	2,116
Father	Senior	.89	.89	.89	.90	.85	.80	.80	.85	.88	.91	.54	.73	.86	.89	.89	2,249
4. Parent's occupation[a]																	
a. Mother	Sophomore	.44	.44	.51	.46	.45	.53	.37	.54	.50	.48	.46	.40	.56	.44	.53	1,921
Mother	Senior	.45	.43	.50	.48	.46	.50	.48	.39	.47	.59	.44	.44	.50	.46	.46	1,970
b. Father	Sophomore	.57	.55	.60	.58	.58	.50	.51	.55	.58	.62	.41	.51	.58	.59	.55	2,220
Father	Senior	.61	.60	.63	.61	.60	.58	.57	.57	.56	.62	.48	.56	.59	.62	.60	2,164
5. Family yearly income	Sophomore	.50	.47	.51	.46	.45	.44	.40	.46	.46	.50	.41	.24	.27	.50	.47	2,652
	Senior	.59	.57	.60	.56	.54	.56	.51	.52	.58	.64	.50	.39	.29	.57	.58	2,650
6. Own house, rent, or other[a]	Sophomore	.56	.54	.57	.57	.55	.52	.52	.56	.60	.57	.59	.55	.52	.57	.53	3,050
	Senior	.58	.58	.58	.57	.60	.56	.55	.58	.63	.59	.59	.55	.60	.58	.56	2,990
7. Race/ethnicity[a]	Sophomore	.78	.79	.78	—	—	—	.73	.80	.78	.83	.78	.77	.79	.76	.80	3,210
	Senior	.79	.79	.80	—	—	—	.76	.82	.80	.80	.83	.75	.80	.80	.79	2,984
8. Home language[a,b]																	
a. Language usually spoken	Sophomore	.74	.67	.73	.81	.53	"e."	.76	.77	.75	.88	.80	.81	.69	.69	.80	639
Language usually spoken	Senior	.78	.80	.76	.81	.83	•	.86	.82	.80	.86	.80	.82	.72	.79	.86	597
b. Other language spoken	Sophomore	.50	.58	.49	.55	.58	•	.55	.54	.64	.60	.51	.53	.64	.32	.64	630
Other language spoken	Senior	.61	.66	.66	.52	.32	•	.48	.68	.60	.74	.60	.61	.66	.61	.64	579
9. Number of siblings																	
a. in college next fall[c]	Senior	.60	.60	.61	.64	.49	.49	.39	.59	.70	.70	.46	.62	.67	.61	.61	2,894
b. in high school next fall	Senior	.85	.85	.86	.88	.80	.75	.79	.83	.89	.91	.82	.86	.90	.85	.88	2,862
10. Mother had job during following periods[d]																	
a. Child was in high school	Sophomore	.70	.67	.72	.72	.66	.[illegible]	.62	.68	.74	.78	.68	.72	.68	.70	—	1,867
	Senior	.72	.70	.74	.75	.67	.61	.62	.75	.82	.74	.68	.75	.75	.72	—	1,756
b. Child was in elementary school	Sophomore	.62	.62	.63	.62	.53	.63	.50	.63	.68	.69	.59	.62	.69	.62	—	1,809
	Senior	.65	.63	.67	.67	.59	.55	.59	.67	.66	.71	.66	.62	.71	.65	—	1,721
c. Before child was in elementary school	Sophomore	.53	.51	.55	.50	.54	.44	.44	.44	.66	.57	.53	.53	.57	.53	—	1,626
	Senior	.53	.47	.56	.54	.47	.37	.48	.50	.61	.50	.53	.52	.51	.53	—	1,613

a. The statistic employed to measure validity is Cramer's V.
b. Including only those students who indicated they had some foreign language background.
c. The parent questionnaire item included non-college postsecondary institutions and was not entirely clear regarding the inclusion of the HS&B student.
d. Includes only those cases where a female answered the parent questionnaire.
e. Sample size is too small (less than 50) to yield sufficiently precise estimates.

BEST COPY

Table A.2. Estimated bias in responses of 1980 sophomores and seniors to family background questionnaire items

Variable	Sophomores		Seniors	
	Child	Parent	Child	Parent
Family yearly income:				
0 to $6,999	7.6	14.1	7.7	14.7
$7,000 to $11,999	12.7	5.9	11.0	6.6
$12,000 to $15,999	19.3	10.8	17.1	10.7
$16,000 to $19,999	18.7	10.8	17.0	10.8
$20,000 to $24,999	17.6	12.1	18.9	12.6
$25,000 to $37,999	12.7	23.3	15.2	21.9
$38,000 and above	11.4	22.9	13.1	22.8
Total	100.0	100.0	100.0	100.0
Mean[a]	$20,600	$24,900	$21,700	$24,600
Own or rent house or apartment:				
Own	77.1	79.9	79.7	82.8
Rent	16.5	17.8	14.8	15.1
Other arrangement	6.4	2.3	5.5	2.1
Total	100.0	100.0	100.0	100.0
Race/ethnicity:				
Hispanic				
Mexican-American	7.5	6.7	6.6	5.6
Cuban	1.2	1.2	1.1	1.2
Puerto Rican	0.9	1.0	1.1	1.1
Other Latin American	2.2	1.3	2.2	1.6
Non-Hispanic				
American-Indian	1.3	2.2	0.8	2.7
Asian	0.8	0.9	1.7	1.6
Black	13.4	14.0	12.5	12.6
White	72.7	72.7	74.0	73.6
Total	100.0	100.0	100.0	100.0
Language usually spoken in home[b]:				
English	58.7	66.5	60.6	63.2
Spanish	31.3	26.9	29.2	27.5
Italian	1.6	0.5	1.2	1.3
Chinese	0.6	1.2	1.0	1.3
French	0.8	0.6	0.5	0.5
German	0.5	0.0	0.3	0.2
Greek	0.0	0.0	0.7	0.7
Portuguese	0.5	0.3	1.2	0.7
Filipino	0.5	0.3	1.0	0.8
Polish	0.3	0.3	0.5	0.7
Other	5.3	3.3	3.8	3.2
Total	100.0	100.0	100.0	100.0
Items in home:				
Specific place for study	47.7	57.3	-	-
Encyclopedia or other references	87.6	88.6	-	-
More than 50 books	84.6	81.3	-	-
Pocket calculator	80.5	86.7	-	-
Daily newspaper	80.0	79.7	-	-
Two or more cars or trucks	78.5	74.6	-	-
Typewriter	75.4	74.1	-	-
Electric dishwasher	54.5	53.8	-	-

Table A.2. (continued). Estimated bias in responses of 1980 sophomores and seniors to family background questionnaire items

Variable	Sophomores		Seniors	
	Child	Parent	Child	Parent
Mother's level of education:				
Less than H.S. graduation	19.8	19.9	21.0	19.5
H.S. graduation only	44.7	35.6	41.4	37.3
Voc., trade, or business school:				
Less than two years	3.7	12.2	4.1	10.7
Two years or more	4.8	2.8	5.2	3.0
College program:				
Less than two years	6.0	12.2	7.2	11.4
Two years	6.9	5.4	6.9	5.2
Four or five years	9.1	8.4	8.8	8.8
Master's degree	3.5	2.6	4.4	3.8
Ph.D., M.D., etc.	1.5	0.8	0.8	0.3
Total	100.0	100.0	100.0	100.0
Father's level of education:				
Less than H.S. graduation	23.3	21.4	24.3	22.7
H.S. graduation only	32.3	24.9	28.8	25.0
Voc., trade, business school:				
Less than two years	3.7	9.1	3.8	6.8
Two years or more	5.8	4.9	5.3	4.0
College program:				
Less than two years	5.3	11.2	6.8	11.9
Two years	6.4	5.6	6.4	4.9
Four or five years	12.2	11.8	12.7	13.5
Master's degree	6.0	6.7	6.8	6.5
Ph.D., M.D., etc.	5.0	4.4	5.2	4.7
Total	100.0	100.0	100.0	100.0
Mother's occupational category:				
Clerical	31.5	34.8	31.1	35.9
Craftsman	2.6	2.1	2.4	2.3
Farmer	0.7	0.2	0.5	0.2
Laborer	2.9	0.9	4.0	0.8
Manager	9.6	10.4	9.3	9.3
Military	0.3	0.1	0.2	0.2
Operative	7.2	10.2	7.7	10.9
Professional 1	11.4	9.6	10.2	9.0
Professional 2	2.3	1.5	2.0	0.8
Protective service	0.3	0.3	0.6	0.6
Sales	6.0	6.9	7.2	5.7
School teacher	9.0	6.1	9.7	7.7
Service	13.6	14.9	13.0	14.6
Technical	2.6	1.9	1.9	2.0
Total	100.0	100.0	100.0	100.0
Father's occupational category:				
Clerical	2.2	5.0	2.2	4.1
Craftsman	16.7	23.3	16.6	22.7
Farmer	4.7	2.1	4.9	2.8
Homemaker	0.3	0.0	0.0	0.0
Laborer	9.8	5.4	9.8	4.9
Manager	20.5	19.3	21.4	19.3
Military	2.3	1.7	1.5	1.3
Operative	13.9	14.7	12.6	13.4
Professional 1	5.7	5.4	8.0	8.0
Professional 2	5.9	5.6	6.5	5.8
Protective service	2.9	2.9	2.9	2.9
Sales	6.0	5.7	5.2	5.8
School teacher	1.9	2.5	2.1	2.0
Service	2.2	3.0	2.3	3.8
Technical	4.9	3.5	3.9	3.2
Total	100.0	100.0	100.0	100.0

Table A.2. (continued). Estimated bias in responses of 1980 sophomores and seniors to family background questionnaire items

Variable	Sophomores		Seniors	
	Child	Parent	Child	Parent
Other language usually spoken in home[b]:				
None	22.1	28.4	21.8	28.0
English	26.7	30.5	25.4	32.0
Spanish	22.9	23.8	22.1	21.6
Italian	3.5	2.7	5.9	2.6
Chinese	1.4	0.3	0.4	0.4
French	6.5	2.4	3.4	1.4
German	7.5	4.8	6.9	3.6
Greek	0.8	0.3	0.9	0.7
Portuguese	1.1	0.6	0.9	0.7
Filipino	0.3	0.2	1.0	1.4
Polish	0.8	0.5	1.9	1.0
Other	6.5	5.6	9.5	6.7
Total	100.0	100.0	100.0	100.0
Siblings in college next fall:[c]				
None	—	—	68.4	59.8
One	—	—	23.5	27.1
Two or more	—	—	8.1	13.0
Total	—	—	100.0	100.0
Siblings in high school next fall:				
None	—	—	47.7	46.6
One	—	—	38.2	38.6
Two or more	—	—	14.1	14.8
Total	—	—	100.0	100.0
Mother usually had a job when child was in high school[d]:				
Did not work	27.2	27.4	25.2	24.5
Worked part-time	22.0	20.4	23.8	25.4
Worked full-time	50.8	52.2	51.0	50.1
Total	100.0	100.0	100.0	100.0
Mother usually had job when child was in elementary school[d]:				
Did not work	34.2	36.4	36.8	36.8
Worked part-time	22.8	22.9	21.5	24.9
Worked full-time	43.0	40.7	41.7	38.3
Total	100.0	100.0	100.0	100.0
Mother usually worked before child went to elementary school[d]:				
Did not work	49.1	56.2	53.2	59.0
Worked part-time	17.2	13.2	13.8	12.3
Worked full-time	33.7	30.6	33.0	28.6
Total	100.0	100.0	100.0	100.0

Note: Details may not add to totals due to rounding.

[a]Means were calculated by using the midpoints of the first six class intervals and a value of $45,000 for the seventh.

[b]Based on only the subset of students who indicated they came from a foreign language background.

[c]The parent questionnaire item included non-college postsecondary institutions and was not entirely clear regarding the inclusion of the HS&B student.

[d]Based on only those cases where a female answered the parent questionnaire.

BEST COPY AVAILABLE

Table A.3. Estimated reliability coefficients for responses of 1980 sophomores and seniors to family background questionnaire items

Variable	Reliability coefficients		No. of observations	
	Sophomores	Seniors	Sophomores	Seniors
Socioeconomic status composite	.82	.86	243	227
Father lives in same household	.84	.86	272	234
Mother lives in same household	.44	.51	272	234
Mother worked when student was in				
High school	.78	.84	248	217
Elementary school	.73	.73	229	208
Before elementary school	.62	.63	201	182
Father's most recent job category	.56	.65	244	219
Father's education	.94	.89	142	160
Years father has lived in U.S.	.70	.77	236	212
Mother's most recent job category	.51	.53	246	225
Mother's education	.85	.90	176	188
Years mother has lived in U.S.	.73	.81	253	225
Student born in U.S.	.96	1.00	269	228
Years student has lived in U.S.	.79	.81	269	230
Student's age	.83	.78	241	216
Racial/ethnic group	.69	.89	266	230
Race	.73	.79	266	230
Religious background	.82	.78	227	206
Family income (one-thirds)	.66	.71	191	181
Family income (one-sevenths)	.69	.75	200	189
Family owns or rents house or apartment	.58	.60	225	210
Number of rooms in home	.64	.71	229	210
Items in home:				
Specific place to study	.33	.30	223	203
Daily newspaper	.57	.60	221	205
Encyclopedia or other reference books	.37	.49	222	205
Typewriter	.70	.75	220	202
Electric dishwasher	.85	.81	215	201
Two or more cars or trucks that run	.72	.70	220	203
More than 50 books	.36	.32	222	204
Own room	.58	.56	224	202
Pocket calculator	.19	.40	221	204
No. of brothers and sisters in college next fall	-	.68	-	210
No. of brothers and sisters in high school next fall	-	.84	-	211

Table A.4. Estimated validity coefficients for responses of 1980 sophomores to 1982 questionnaire items about grades, coursework, and courses, by selected student characteristics*

Variable	All students	Sex		Race/ethnicity			Test score quartile				SES composite		
		Male	Female	White	Hispanic	Black	Low	2	3	High	Low	Med.	High
Grades in high school	.77	.76	.79	.80	.66	.65	.54	.66	.75	.82	.68	.77	.82
Coursework, 10th grade through end of 1981-82 school year (semesters):													
Mathematics	.66	.65	.67	.71	.55	.52	.41	.60	.70	.78	.56	.66	.66
English or literature	.28	.24	.31	.34	.25	.26	.20	.30	.35	.33	.29	.29	.29
French	.87	.84	.88	.92	.82	.80	.79	.81	.84	.89	.85	.87	.87
German	.87	.85	.89	.89	.49	.86	.65	.78	.91	.91	.75	.88	.86
Spanish	.86	.83	.88	.88	.80	.83	.79	.86	.85	.88	.82	.85	.85
History or social studies	.39	.39	.39	.40	.29	.31	.31	.36	.43	.49	.32	.40	.41
Science	.70	.66	.73	.76	.58	.61	.45	.58	.70	.80	.60	.70	.70
Courses taken through 1981-82 school year (yes/no):													
Second-year algebra	.68	.66	.69	.69	.60	.56	.45	.56	.66	.61	.63	.68	.62
Geometry	.85	.84	.85	.85	.81	.81	.80	.81	.84	.74	.82	.86	.76
Trigonometry	.63	.63	.64	.63	.67	.60	.27	.55	.64	.54	.61	.63	.58
Calculus	.67	.64	.70	.67	.55	.61	.00	.45	.52	.68	.57	.61	.71
Physics	.80	.81	.78	.86	.54	.59	.29	.55	.82	.93	.55	.81	.90
Chemistry	.87	.85	.89	.88	.79	.81	.66	.81	.88	.88	.81	.87	.87

*The statistics in this table are based on comparisons of student responses with high school transcript data.

Table A.5. Estimated bias in responses of 1980 sophomores to 1982 questionnaire items about grades and amount of coursework in selected areas, by selected student characteristics

Variable and data source	All students	Sex		Race/ethnicity			Test score quartile				SES composite		
		Male	Female	White	Hispanic	Black	1	2	3	4	Low	Med.	High
Grade average (letters—A = 4, B = 3, C = 2, D = 1)													
Student	2.84	2.71	2.96	2.91	2.57	2.62	2.45	2.60	2.90	3.31	2.64	2.85	3.07
Transcript	2.62	2.51	2.73	2.71	2.39	2.31	2.21	2.41	2.68	3.08	2.44	2.63	2.84
Difference (bias)	.22	.20	.23	.20	.18	.31	.24	.19	.22	.23	.20	.22	.23
Mathematics (no. of semesters, grade 10 through end of 1981-82 school year):													
Student	4.15	4.31	4.02	4.15	3.97	4.50	3.63	3.65	4.08	5.02	3.68	4.07	4.76
Transcript	3.07	3.17	3.03	3.27	2.39	2.65	1.79	2.46	3.29	4.44	2.27	3.03	4.02
Difference (bias)	1.08	1.14	.99	.88	1.58	1.85	1.84	1.19	.79	.58	1.41	1.04	.74
English or literature (no. of semesters, grade 10 through end of 1981-82 school year):													
Student	5.81	5.75	5.89	5.82	5.73	5.98	5.74	5.78	5.84	5.98	5.72	5.82	5.97
Transcript	5.56	5.48	5.67	5.58	5.47	5.63	5.47	5.55	5.58	5.78	5.48	5.56	5.72
Difference (bias)	.25	.27	.22	.24	.26	.35	.27	.23	.26	.20	.24	.26	.25
History or social studies (no. of semesters, grade 10 through end of 1981-82 school year):													
Student	4.58	4.52	4.66	4.62	4.43	4.58	4.38	4.55	4.60	4.72	4.49	4.53	4.69
Transcript	4.63	4.60	4.70	4.71	4.45	4.44	4.32	4.61	4.71	4.88	4.51	4.64	4.81
Difference (bias)	-.05	-.08	-.04	-.09	-.02	.14	.06	-.06	-.11	-.16	-.02	-.11	-.12
Science (no. of semesters, grade 10 through end of 1981-82 school year):													
Student	3.43	3.58	3.30	3.47	3.13	3.46	2.79	2.87	3.37	4.38	2.92	3.26	4.09
Transcript	2.87	2.99	2.78	3.00	2.33	2.59	1.95	2.27	2.90	4.10	2.29	2.78	3.66
Difference (bias)	.56	.59	.52	.47	.80	.87	.84	.60	.47	.28	.63	.48	.43

BEST COPY AVAILABLE

Table A.6. Estimated reliability coefficients for responses of 1980 sophomores and seniors to school-related questionnaire items

Variable	Reliability coefficients		No. of observations	
	Sophomores	Seniors	Sophomores	Seniors
Instruction methods used in course taken this year (used never, seldom, fairly often, or frequently):				
Listening to the teacher's lecture	—	.25	—	229
Participating in student-centered discussions	—	.28	—	220
Working on a project or in a laboratory	—	.30	—	221
Writing essays, themes, or stories	—	.33	—	219
Having individualized instruction	—	.19	—	224
Using teaching machines or computer-assisted instruction	—	.05	—	229
Extent to which disciplinary matters are problems in your school (often happens, sometimes happens, never happens):				
Students don't attend school	.36	—	256	—
Students cut classes, even if they attend school	.38	—	265	—
Students talk back to teachers	.23	—	263	—
Students refuse to obey instructions	.19	—	263	—
Students get in fights with each other	.30	—	261	—
Students attack or threaten to attack teachers	.18	—	263	—
Rules that are enforced in your school:				
School grounds closed to students at lunch time	.38	—	276	—
Students responsible to the school for property damage	.26	—	276	—
Hall passes required	.43	—	276	—
"No smoking" rules	.36	—	276	—
Rules about student dress	.54	—	276	—
School ratings (poor, fair, don't know, good, excellent):				
Condition of buildings and classrooms	.39	.52	259	225
Library facilities	.29	.39	257	224
Quality of academic instruction	.20	.32	257	219
Reputation in the community	.23	.44	252	221
Teacher interest in students	.23	.35	254	221
Effective discipline	.21	.26	252	221
Fairness of discipline	.21	.24	253	221
School spirit	.31	.38	257	221
High school has minimum competency or proficiency test—a special test that must be passed to get a H.S. diploma (yes, don't know, no)	.45	.68	256	225
How much has each of the following interfered with your education at this school? (not at all, somewhat, a great deal):				
Courses are too hard	—	.23	—	225
Find it hard to adjust to school routine	—	.14	—	223
Poor teaching	—	.24	—	222
Poor study habits	—	.18	—	225
Courses are too easy	—	.22	—	223
How many of the students in your class were black? (none, few, about half, most, all):				
In first grade	.76	.71	222	204
In sixth grade	.68	.81	220	199
In ninth grade	.75	.83	221	198
How many of the students in your class were Hispanic? (none, few, about half, most, all):				
In first grade	.58	.70	213	201
In sixth grade	.63	.68	209	195
In ninth grade	.67	.75	211	193
No. of school changes due to family moving	.77	—	263	—
When began attending this school	—	.89	—	230
Kindergarten attendance	.93	—	272	—

Table A.7. Estimates of coefficients of consistency between responses of 1980 sophomores and seniors and those of their parents to questionnaire items about high school grades, expectations, and attitudes, by selected student characteristics and sex of parent questionnaire respondent

Variable	Cohort	All students	Sex		Race/ethnicity			Test score quartile				SES composite			Respondent		n
			Male	Female	White	Hispanic	Black	1	2	3	4	Low	Med.	High	Mother	Father	
1. High school grade average	Sophomore	.75	.75	.77	.79	.67	.62	.59	.68	.75	.78	.69	.77	.81	.74	.80	3,524
	Senior	.76	.73	.76	.77	.69	.70	.56	.70	.74	.79	.69	.76	.82	.74	.77	3,234
2. Parental influence on post-high school plans																	
a. Mother's influence	Senior	.18	.18	.18	.18	.23	.19	.17	.18	.19	.15	.16	.18	.15	.19	.18	2,819
b. Father's influence	Senior	.21	.18	.24	.22	.19	.21	.14	.24	.18	.25	.18	.20	.16	.24	.18	2,353
3. Age student expects to																	
a. Marry	Sophomore	.42	.32	.47	.44	.39	.31	.33	.44	.40	.44	.42	.42	.34	.42	.41	2,278
	Senior	.57	.45	.60	.58	.58	.40	.54	.58	.52	.53	.62	.52	.49	.56	.55	2,118
b. Have first child	Sophomore	.39	.28	.46	.40	.23	.52	.30	.34	.36	.43	.38	.37	.28	.40	.38	2,038
	Senior	.51	.36	.58	.50	.53	.51	.51	.47	.48	.44	.59	.44	.38	.50	.50	1,797
c. Start regular job	Sophomore	.32	.36	.29	.34	.28	.21	.18	.30	.25	.33	.22	.32	.33	.32	.33	2,753
	Senior	.48	.47	.46	.51	.46	.36	.34	.46	.42	.45	.38	.46	.44	.47	.49	2,617
d. Live in own home	Sophomore	.27	.21	.34	.27	.33	.12	.30	.29	.30	.19	.31	.28	.21	.29	.25	2,529
	Senior	.40	.28	.50	.43	.46	.23	.37	.41	.36	.42	.40	.40	.29	.38	.43	2,373
e. Finish full-time education	Sophomore	.43	.44	.43	.47	.39	.25	.29	.44	.41	.46	.36	.46	.36	.44	.41	2,385
	Senior	.48	.46	.49	.50	.44	.37	.44	.42	.45	.49	.47	.46	.40	.47	.49	2,335
4. Occupational expectations, age 30[a]	Sophomore	.31	.34	.27	.32	.38	.31	.30	.34	.34	.33	.32	.32	.36	.31	.34	3,041
	Senior	.37	.34	.38	.40	.38	.37	.35	.38	.40	.43	.36	.39	.39	.37	.39	2,968
5. College expectations when child was in																	
a. Grade 6 or 7	Sophomore	.43	.44	.42	.46	.38	.25	.36	.36	.32	.29	.41	.38	.23	.44	.41	2,188
b. Grade 8 or 9	Sophomore	.45	.46	.43	.49	.31	.19	.41	.38	.38	.41	.43	.44	.31	.46	.42	2,625
Grade 8 or 9	Senior	.40	.40	.40	.44	.35	.21	.31	.30	.36	.30	.34	.38	.27	.40	.41	2,212
c. Grade 10	Senior	.42	.46	.39	.47	.28	.13	.29	.40	.42	.28	.41	.41	.23	.39	.45	2,203
d. Grade 11	Senior	.46	.51	.41	.50	.32	.30	.41	.46	.46	.29	.43	.49	.25	.44	.50	2,404
6. Mother's aspirations for child's education[b]	Sophomore	.57	.59	.56	.60	.51	.45	.47	.54	.51	.51	.49	.58	.47	.57	—	1,595
	Senior	.59	.60	.59	.64	.52	.40	.46	.55	.61	.60	.58	.52	.56	.59	—	1,556
7. Student's ability to complete college	Sophomore	.40	.42	.37	.44	.31	.23	.30	.29	.30	.34	.38	.36	.32	.39	.41	2,983
	Senior	.42	.44	.39	.44	.34	.32	.31	.40	.24	.18	.39	.41	.34	.45	.45	2,983
8. Sex role attitudes																	
a. Working mother is good mother	Sophomore	.18	.16	.19	.18	.18	.05	.13	.14	.19	.21	.17	.16	.20	.19	.12	3,140
b. Male should be the achiever	Sophomore	.21	.20	.20	.25	.19	.08	.14	.18	.25	.26	.19	.19	.27	.24	.16	3,097
c. Women happiest making a home	Sophomore	.18	.15	.19	.17	.18	.12	.09	.13	.12	.21	.10	.14	.22	.18	.16	3,027

[a]The statistic employed to measure consistency is Cramer's V.
[b]Includes only those cases where the mother answered the Parent Questionnaire.

BEST COPY

Table A.8. Relative bias between parent and child (1980 sophomores and seniors) responses to questionnaire items about grades, plans, and attitudes

Variable	Sophomores		Seniors	
	Child	Parent	Child	Parent
Grades in high school:				
Mostly A's (90-100)	10.8	12.9	12.5	16.4
A- or B+ (85-89)	18.1	20.2	22.5	23.7
Mostly B's (80-84)	17.9	17.0	21.7	19.2
B- or C+ (75-79)	23.2	24.7	25.5	24.2
Mostly C's (70-74)	13.7	13.8	12.6	10.7
C or D+ (65-69)	8.6	8.2	4.6	5.0
Mostly D's (60-64)	1.9	2.2	0.6	0.7
Mostly Below D	0.8	1.0	0.0	0.0
Total	100.0	100.0	100.0	100.0
Mother's influence on child's plans after high school:				
Not at all	—	—	11.4	13.8
Somewhat	—	—	44.6	41.8
A great deal	—	—	44.1	44.4
Total	—	—	100.0	100.0
Father's influence on child's plans after high school:				
Not at all	—	—	14.1	16.0
Somewhat	—	—	45.8	46.1
A great deal	—	—	40.1	37.9
Total	—	—	100.0	100.0
Age child expects (is expected) to get married:				
Have already done this	0.3	0.6	1.3	3.7
20 or under	21.7	15.8	19.2	13.7
21 to 24	50.6	50.5	51.5	51.0
25 or more	27.4	32.9	28.0	31.6
Total	100.0	100.0	100.0	100.0
Age child expects (is expected) to have first child:				
Have already done this	0.6	0.8	1.8	2.1
20 or under	7.4	5.8	6.3	5.3
21 to 24	43.4	37.6	37.6	35.7
25 or more	48.6	55.8	54.3	56.9
Total	100.0	100.0	100.0	100.0
Age child expects (is expected) to have first regular job:				
Have already done this	9.3	3.6	10.0	12.6
19 or under	56.8	35.3	36.1	22.3
20 to 22	20.3	40.0	31.4	41.4
23 or more	13.6	21.1	16.5	23.7
Total	100.0	100.0	100.0	100.0
Age child expects (is expected) to live in own home or apartment:				
Have already done this	0.6	0.8	2.2	5.4
20 or under	66.1	37.8	49.0	29.4
21 to 24	27.5	50.0	42.8	54.0
25 or more	5.8	11.4	6.0	11.2
Total	100.0	100.0	100.0	100.0
Age child expects (is expected) to finish fulltime education:				
Have already done this	17.9	0.1	1.4	5.2
20 or under	54.0	28.7	31.6	18.1
21 to 24	23.9	60.6	57.0	64.4
25 or more	4.2	10.6	10.0	12.3
Total	100.0	100.0	100.0	100.0

Table A.8. (continued) Relative bias between parent and child (1980 sophomores and seniors) responses to questionnaire items about grades, plans, and attitudes

Variable	Sophomores		Seniors	
	Child	Parent	Child	Parent
Child's expected occupation when 30 years old:				
Clerical	9.1	11.2	10.6	10.8
Craftsman	8.6	10.0	7.1	7.2
Farmer	2.8	1.6	1.6	1.2
Homemaker	5.3	5.6	2.9	5.8
Laborer	2.1	1.9	2.0	2.1
Manager	3.8	5.6	7.1	9.1
Military	3.4	2.9	2.2	2.3
Operative	3.1	2.7	2.9	2.7
Professional, advanced degree	24.4	24.8	26.5	25.2
Professional, other	13.2	10.4	13.3	9.9
Proprietor or owner	2.9	2.6	3.5	3.3
Protective service	1.8	1.4	1.8	1.7
Sales	1.9	1.9	1.9	3.2
School teacher, elementary or sec.	2.7	4.5	3.8	4.1
Service	4.2	3.9	3.6	2.6
Technical	8.2	8.5	8.1	8.3
Not working	2.6	0.5	1.2	0.5
Total	100.0	100.0	100.0	100.0
Child expected to go to college when in earlier grades (pct "yes")				
6th or 7th grade	66.4	65.1	—	—
8th or 9th grade	74.7	69.0	71.4	67.5
10th grade	—	—	73.6	67.9
11th grade	—	—	77.0	71.3
Mother's educational desires for child:[a]				
Less than H.S. graduation	0.5	0.4	0.6	0.1
High school graduation only	16.4	7.5	11.7	5.9
Vocational, trade, or business school:				
Less than two years	3.8	10.4	4.6	11.8
Two years or more	9.2	9.6	11.4	8.8
College program				
Less than two years	1.6	8.7	2.4	7.1
Two or more years	11.0	8.3	11.0	8.6
Finish 4- or 5-year program	32.5	34.4	33.9	34.6
Master's degree	9.2	10.4	9.5	13.0
Ph.D., M.D., or equivalent	15.7	10.4	14.7	10.1
Total	100.0	100.0	100.0	100.0
Ability to complete college:				
Yes, definitely	39.0	58.0	49.[illegible]	64.1
Yes, probably	34.8	30.4	32.8	27.2
No, probably not (not sure)	22.9	8.6	16.0	6.5
No, definitely not	3.3	2.9	1.9	2.2
Total	100.0	100.0	100.0	100.0
A working mother of pre-school children can be just as good a mother as the woman who doesn't work:				
Agree strongly	19.0	23.1	—	—
Agree	44.2	31.8	—	—
Disagree	29.0	30.2	—	—
Disagree strongly	7.8	14.9	—	—
Total	100.0	100.0	—	
It is usually better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and family:				
Agree strongly	11.7	20.0	—	—
Agree	38.1	34.8	—	—
Disagree	38.0	34.9	—	—
Disagree strongly	12.3	10.3	—	—
Total	100.0	100.0	—	—
Most women are happiest when they are making a home and caring for children:				
Agree strongly	10.8	11.7	—	—
Agree	38.0	28.7	—	—
Disagree	43.4	47.8	—	—
Disagree strongly	7.8	11.8	—	—
Total	100.0	100.0	—	—

Note: Details may not add to totals due to rounding.

[a]Based on only those cases where the mother answered the parent questionnaire.

Table A.9. Estimates of coefficients of consistency between responses of 1980 sophomores and seniors and those of their parents to questionnaire items about financial matters and college choice factors, by selected student characteristics and sex of parent questionnaire respondent

Variable	Cohort	All students	Sex		Race/ethnicity			Test score quartile				SES composite			Respondent		n
			Male	Female	White	Hispanic	Black	1	2	3	4	Low	Med.	High	Mother	Father	
1. Schooling expenses, Sept '80—Aug. '81	Senior	.61	.60	.61	.66	.41	.41	.43	.53	.58	.62	.56	.55	.59	.57	.64	2,789
2. Living expenses, Sept. '80—Aug. '81	Senior	.22	.20	.24	.25	.11	.11	.07	.22	.30	.33	.17	.21	.33	.20	.26	2,790
3. Payment for living and school expenses from																	
a. Parents	Senior	.47	.42	.51	.50	.20	.43	.25	.37	.46	.57	.23	.38	.47	.45	.47	2,428
b. Spouse	Senior	.12	.04	.18	.14	.22	.01	.02	.26	.16	-.01	.16	.13	.04	.08	.17	1,924
c. Other relative	Senior	.16	.18	.15	.19	.06	.09	.12	.25	.28	.13	.09	.17	.21	.11	.28	1,964
d. Summer earnings ('80)	Senior	.18	.16	.18	.20	.16	-.01	.12	.13	.18	.29	.11	.19	.26	.14	.26	2,243
e. Earnings Sept. 80—Aug. '81	Senior	.27	.26	.27	.29	.17	.12	.13	.27	.30	.32	.18	.28	.34	.22	.33	2,021
f. Savings	Senior	.17	.20	.17	.19	.13	.20	.11	.15	.15	.32	.13	.15	.28	.17	.20	1,958
g. State scholarship or grant	Senior	.33	.26	.39	.36	.23	.17	.23	.29	.39	.39	.36	.30	.29	.31	.35	2,105
h. Federal scholarship or grant	Senior	.33	.32	.36	.32	.35	.29	.30	.31	.39	.33	.35	.32	.30	.35	.27	2,012
i. Other scholarship or grant	Senior	.38	.39	.37	.42	.38	.21	.17	.25	.36	.48	.26	.40	.48	.34	.43	2,046
j. State loan	Senior	.17	.12	.23	.22	.04	.05	.05	-.01	.28	.27	.08	.18	.23	.12	.25	2,010
k. Federal loan	Senior	.16	.18	.14	.20	.18	.02	.05	.13	.20	.22	.09	.18	.20	.16	.17	2,020
l. Other loan	Senior	.07	.06	.09	.08	-.05	.02	-.02	.60	.08	.20	.01	.10	.11	.02	.15	2,004
m. Social security or VA benefits	Senior	.47	.39	.58	.53	.19	.34	.24	.47	.57	.70	.39	.51	.56	.45	.47	2,119
Mean		(.25)	(.23)	(.28)	(.28)	(.17)	(.15)	(.14)	(.22)	(.29)	(.32)	(.19)	(.25)	(.29)	(.22)	(.29)	—
4. Estimated cost of schooling*																	
a. Public junior or community college	Sophomore	.14	.12	.16	.17	.45	.03	.14	.04	.12	.25	.04	.16	.20	.15	.13	1,665
	Senior	.25	.25	.25	.24	.25	.35	.29	.22	.33	.25	.20	.25	.31	.23	.29	1,962
b. State 4-yr college or university	Sophomore	.12	.08	.17	.15	.05	.10	.15	.10	.18	.14	.18	.12	.14	.14	.11	1,627
	Senior	.15	.13	.18	.14	.14	.24	.15	.11	.22	.14	.18	.14	.14	.13	.18	1,890
c. Private 4-yr college or university	Sophomore	.12	.10	.14	.14	.02	.07	.21	.01	.03	.20	.14	.06	.20	.14	.08	1,337
	Senior	.06	.08	.04	.06	-.02	.15	.01	.02	.10	.10	.06	.03	.09	.04	.10	1,600
Mean		(.14)	(.13)	(.16)	(.15)	(.15)	(.16)	(.14)	(.08)	(.16)	(.18)	(.13)	(.13)	(.18)	(.14)	(.15)	—
5. Factors important in choosing a college*																	
a. College expenses	Senior	.24	.19	.27	.25	.04	.13	.04	.14	.26	.30	.13	.16	.28	.23	.25	1,702
b. Availability of financial aid	Senior	.44	.35	.50	.43	.24	.25	.25	.36	.43	.50	.26	.38	.34	.44	.42	1,699
c. Availability of specific courses	Senior	.11	.12	.11	.12	.01	.17	.04	.11	.12	.12	.02	.10	.19	.11	.12	1,676
d. Reputation in academic areas	Senior	.15	.14	.16	.15	.18	.06	.10	.15	.13	.12	.17	.11	.14	.16	.14	1,696
e. Reputation in athletic program	Senior	.32	.37	.25	.33	.32	.26	.35	.23	.24	.34	.32	.35	.31	.31	.32	1,686
f. Social life of college	Senior	.18	.19	.17	.16	.26	.08	.17	.13	.19	.17	.25	.14	.19	.18	.18	1,681
g. Able to live at home and attend college	Senior	.56	.50	.60	.58	.32	.45	.37	.43	.58	.60	.47	.53	.60	.55	.59	1,698
Mean		(.29)	(.27)	(.29)	(.29)	(.20)	(.20)	(.19)	(.22)	(.28)	(.31)	(.23)	(.25)	(.29)	(.28)	(.29)	—

*Includes only those cases where the student thought he or she might attend a postsecondary institution and the student's parent said the student was.

Table A.10. Relative bias between parent and child (1980 sophomores and seniors) responses to questionnaire items about financial matters and college choice factors.

Variables	Sophomores		Seniors	
	Child	Parent	Child	Parent
Money required by child for schooling expenses, Sept. 1980 - Aug. 1981:				
None	—	—	23.2	27.7
Less than $500	—	—	14.4	10.8
$500 to $999	—	—	14.0	17.5
$1,000 to $1,999	—	—	17.2	18.1
$2,000 to $4,000	—	—	20.2	16.7
More than $4,000	—	—	11.0	9.1
Total	—	—	100.0	100.0
Money required by child for living expenses next year:				
None, live at home	—	—	22.6	31.1
None, other reasons	—	—	4.1	10.4
Less than $1,000	—	—	15.8	9.7
$1,000 to $1,999	—	—	20.1	15.5
$2,000 to $2,999	—	—	14.4	16.6
$3,000 to $3,999	—	—	9.4	8.8
$4,000 to $4,999	—	—	5.8	4.1
$5,000 to $5,999	—	—	4.3	2.8
$6,000 to $10,000	—	—	1.6	0.8
More than $10,000	—	—	1.8	0.3
Total	—	—	100.0	100.0
To pay for living and school expenses from Sept. 1980 through Aug. 1981, child to receive money from following sources (pct. reporting some money from the sources):				
Parents	—	—	77.2	74.8
Spouse	—	—	10.2	7.8
Other relatives	—	—	20.6	8.0
Summer earnings, 1980	—	—	86.7	62.2
Earnings, Sept. 1980-Aug. 1981	—	—	73.8	63.2
Savings	—	—	71.0	35.0
State scholarship or grant	—	—	33.0	20.0
Federal scholarship or grant	—	—	26.0	17.0
Other scholarship or grant	—	—	29.0	18.0
State loan	—	—	17.0	8.0
Federal loan	—	—	15.0	11.0
Other loan	—	—	16.0	6.0
Social Security or VA benefits	—	—	13.0	10.0
Estimated schooling expenses at a public junior or community college:				
Under $500	14.6	19.2	18.1	18.7
$500 to $1,000	35.1	39.8	34.8	39.6
$1,001 to $2,000	34.4	27.7	29.6	28.0
$2,001 to $3,000	11.7	9.8	12.5	9.8
$3,001 to $5,000	3.5	2.9	4.4	3.3
$5,001 to $7,000	0.7	0.6	0.7	0.5
Total	100.0	100.0	100.0	100.0
Estimated schooling expenses at a state 4-year college or university:				
Under $500	1.7	1.6	1.3	1.9
$500 to $1,000	7.3	12.4	9.3	11.4
$1,001 to $2,000	21.1	28.2	20.5	28.3
$2,001 to $3,000	29.5	27.5	28.9	29.2
$3,001 to $5,000	30.4	21.9	31.8	21.4
$5,001 to $7,000	10.0	8.4	8.2	7.7
Total	100.0	100.0	100.0	100.0
Estimated schooling expenses at a private 4-year college or university:				
Under $500	1.3	0.5	0.8	0.4
$500 to $1,000	2.3	2.0	2.0	1.5
$1,001 to $2,000	5.8	5.2	5.8	6.0
$2,001 to $3,000	14.4	16.7	13.6	13.8
$3,001 to $5,000	25.8	31.8	30.8	32.9
$5,001 to $7,000	50.4	43.8	47.1	45.4
Total	100.0	100.0	100.0	100.0
Factors important in choosing a college (pct. "very important")				
College expenses	—	—	36.6	52.5
Availability of financial aid	—	—	40.2	43.6
Availability of specific courses	—	—	68.7	69.9
Reputation in academic areas	—	—	52.8	75.2
Reputation in athletic programs	—	—	10.7	9.2
Social life of the college	—	—	28.1	11.0
Able to live at home and attend college	—	—	21.7	32.6

Table A.11. Approximate standard errors of correlation coefficients*

Subgroup	Sophomores			Seniors		
	r=0	r=.50	r=.90	r=0	r=.50	r=.90
Student-transcript comparisons:						
Total population	–	–	–	.019	.015	.004
Male or female students	–	–	–	.028	.021	.005
Test or SES quartile subgroups	–	–	–	.042	.032	.008
White students	–	–	–	.025	.019	.005
Black students	–	–	–	.053	.043	.010
Hispanic students	–	–	–	.043	.032	.008
Parent-child comparisons:						
Total population	.027	.020	.005	.028	.021	.005
Male or female students	.040	.030	.008	.040	.030	.008
Test or SES quartile subgroups	.055	.041	.011	.055	.041	.011
White students	.032	.024	.006	.033	.025	.006
Black students	.072	.054	.014	.079	.059	.015
Hispanic students	.084	.063	.016	.091	.068	.017
Female respondents (Parent Questionnaire)	.034	.026	.006	.036	.027	.007
Male respondents (Parent Questionnaire)	.045	.034	.009	.045	.034	.009
Twin/co-twin comparisons:						
Total population	.094	.071	.018	.102	.077	.015

*The standard errors presented in the table were calculated from the formula $var(r) \approx D(1 - r^2)^2/(.8n)$, where r is the estimated correlation coefficient, D is a survey design adjustment factor, and n is the sample size (Kendall & Stewart, 1958). Conservative values of D were employed: 2.0 for parent-child and twin/co-twin comparisons and 4.0 for comparisons based on transcript data (Tourangeau *et al.*, in preparation). Sample sizes (Chapter 1) were reduced by 20 percent to adjust for cases not usable in the analysis due to item non-response, "don't know" replies, etc.

The analysis for a few items was restricted to subsets of cases, viz., language in home, mother had job, mother's aspirations for child's education, estimated cost of school, and college choice factors. For the first item, the tabled standard errors should be doubled; for the remaining ones, they should be increased by 30 percent.

To illustrate the use of table A.11, consider the estimated correlation coefficient of .87 between reports of father's education level by sophomores and their parents (table 3.1). The standard error is found by linear interpolation between the values for r = .50 and r = .90 in table A.11 to be approximately .005.

Table A.12. Approximate standard errors of bias estimates (percentage points)—parent as standard*

Correlation coefficient	Percentage (P)		
	P = 50	P = 20 (or 80)	P = 10 (or 90)
0.2	2.0	1.6	1.2
0.4	1.7	1.4	1.0
0.6	1.4	1.1	0.8
0.8	1.0	0.8	0.6

*The size of the standard error is a function of the correlation (r) between child and parent responses and the percentage values for child (P_1) and parent (P_2). Computation of standard errors were based on the equation $Var(b) = Var(P_1) + Var(P_2) - 2cov(P_1,P_2)$, where $b = P_1 - P_2$. When $P_1 = P_2 = P$, $Var(b) = 2(1 - r)Var(P)$. Estimates were adjusted for item nonresponse, "don't know" responses, etc., by reducing n by 20 percent and for survey design effect by increasing the simple random sampling estimates by 40 percent. The standard error estimates may be used for either cohort.

The analysis for a few items was restricted to subsets of cases, viz., language in home, mother had job, mother's aspirations for child's education, estimated cost of school, and college choice factors. For the first item, the tabled standard errors should be doubled; for the remaining ones, they should be increased by 30 percent.

To illustrate the use of Table A.12, consider the estimated bias in the variable "home has 50 or more books." The estimated bias is 4 percentage points since 85 percent of the parents and 81 percent of their children replied "yes;" and the coefficient of correlation between child and parent responses is .35. (See tables A.2 and A.3) Thus the standard error of the bias estimate from table A.12 is about 1.4 percentage points and the bias differs from zero by 2.9 standard errors (4/1.4 = 2.9).

Table A.13. Approximate standard errors of bias estimates—transcript as standard[a]

Subgroup	Grade average (letters)	Amount of coursework (semesters)
All students	.10	.028
Males, females	.12	.035
Test quartile	.15	.052
SES quartile	.16	.050
White students	.11	.032
Black students	.25	.059
Hispanic students	.20	.062

[a]"Balanced repeated replicates" estimates of standard errors were calculated for each bias value. This table shows the mean of the standard error values for the four coursework areas.

☆U.S. GOVERNMENT PRINTING OFFICE: 1984 421 054 4241

NCES 84-216